A Bond-Energy/Bond-Order and Populations Relationship

Barbaro Zulueta,[†] Sonia V. Tulyani,[‡] Phillip R. Westmoreland,[¶] Michael J. Frisch,[§] E. James Petersson,[‖] George A. Petersson,[⊥,#] and John A. Keith[*,†]

[†]*Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, United States*

[‡]*Formerly Chemical Engineering Department, University of Massachusetts Amherst,618 N. Pleasant Street, Amherst, MA 01003*

[¶]*Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina 27695, United States*

[§]*Gaussian, Inc., Wallingford, Connecticut 06492, United States*

[‖]*Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States*

[⊥]*Institute for Computational Molecular Science, Temple University, Philadelphia, Pennsylvania 19122, United States*

[#]*Formerly Hall-Atwater Laboratories of Chemistry,Wesleyan University, Middletown, Connecticut 06459*

E-mail: jakeith@pitt.edu

Abstract

We report an analytical Bond Energy from Bond Orders and Populations (BEBOP) model that provides intramolecular bond energy decompositions for chemical insight into the thermochemistry of molecules. The implementation reported here employs

a minimum basis set Mulliken population analysis on well-conditioned Hartree-Fock orbitals to decompose total electronic energies into physically interpretable contributions. The model's parameterization scheme is based on atom-specific parameters for hybridization and atom pair-specific parameters for short-range repulsion and extended Hückel-type bond energy term fitted to reproduce CBS-QB3 thermochemistry data. The current implementation is suitable for molecules involving H, Li, Be, B, C, N, O, and F atoms, and it can be used to analyze intramolecular bond energies of molecular structures at optimized stationary points found from other computational methods. This first-generation model brings the computational cost of a Hartree-Fock calculation using a large triple-zeta basis set, and its atomization energies are comparable to those from widely used hybrid Kohn-Sham density functional theory (DFT, as benchmarked to 109 species from the G2/97 test set and an additional 83 reference species). This model should be useful for the community by interpreting overall *ab initio* molecular energies in terms of physically insightful bond energy contributions, e.g. bond dissociation energies, resonance energies, molecular strain energies, and qualitative energetic contributions to the activation barrier in chemical reaction mechanisms. This work reports a critical benchmarking of this method as well as discussions of its strengths and weaknesses compared to hybrid DFT (i.e., B3LYP, M062X, PBE0, and APF methods), and other cost-effective approximate Hamiltonian semi-empirical quantum methods (i.e., AM1, PM6, PM7, and DFTB3).

Introduction

The qualitative relation of bond energies to bond orders and hybridization is among the most fundamental concepts of chemistry.[1–3] It is therefore surprising that relatively few modern computational quantum chemistry methods explicitly use these concepts. For example, the natural bond order analysis (NBO) method of Weinhold[4–6] transforms a variationally optimized quantum mechanical wavefunction into a localized form that corresponds to one-center lone pairs and two-center bonds that conveniently are related to chemical Lewis structures.

Alternatively, reactive forcefield methods of Brenner,[7] Pettifor,[8] Tersoff,[9] and van Duin and Goddard[10] begin from the *ansatz* that bond energies can be realized from *a priori* bond orders approximated from fitted mathematical expressions.

To our knowledge, our BEBOP model is the first computational approach that starts from a minimum basis set Mulliken orbital population analysis of HF calculations and then uses those to construct total system energies (including approximate correlation and zero point energies implicitly) in terms of a variety of insightful chemical bonding and anti-bonding energy contributions. The net product is a "black-box" method that only requires the cost of HF theory while giving the accuracy of higher level quantum chemistry methods. In this way BEBOP relates to Miller and Manby's OrbNet Denali method that uses machine learning (ML) to map symmetry-adapted atomic orbital features from relatively low-cost semi-empirical quantum mechanics (SQM) theory to energies of high-level DFT theory calculations (i.e., ωB97X-D3/def2-TZVP).[11] Key differences with the BEBOP approach are that much more robust (and much more computationally expensive) HF orbitals are used as a starting point, these are mapped to more accurate energies from higher-level *ab initio* thermochemical methods, and the system energies are partitioned into useful chemical bonding quantities that go beyond what is done in NBO and energy decomposition analysis (EDA)[12–16] methods. At the present time, we only recommend its use for studies of thermochemical analyses of molecules whose geometries have been optimized to stationary points from other computational methods. As explained below, the BEBOP method uses a parametrization scheme that requires fewer parameters than most SQM and reactive forcefield methods and thus indicates promise for eventual transferability across large parts of the periodic table. There are also clear deficiencies in the present first-generation model that can be improved upon, though we note that many hybrid density functional methods also suffer from similar deficiencies and still have been widely adopted.

Theory: Bond Energy and Bond Order

Within the framework of an independent particle model such as HF or DFT, the total electronic energy can be expressed as a sum of one-electron energies. The relationship between bond energy and bond order follows from the energy of an electron in a molecular orbital, φ:

$$E = \frac{\left\langle \varphi \middle| \hat{H} \middle| \varphi \right\rangle}{\langle \varphi | \varphi \rangle}, \tag{1}$$

where $\hat{H}$ is the effective one-electron Hamiltonian. For simplicity we assume that for a molecular orbital(MO), φ^{MO}, is a linear combination of just two atomic orbitals (AO), χ_1^{AO} and χ_2^{AO}:

$$\varphi^{MO} = c_1\chi_1^{AO} + c_2\chi_2^{AO}. \tag{2}$$

Substitution of Eq.(2) in Eq.(1) gives:

$$E = \frac{c_1^2 H_{11} + 2c_1c_2H_{12} + c_2^2H_{22}}{c_1^2 + 2c_1c_2S_{12} + c_2^2} = c_1^2 H_{11} + 2c_1c_2H_{12} + c_2^2H_{22}, \tag{3}$$

where S_{12} is the AO overlap integral, and we have assumed normalization of the molecular orbital: $\langle \varphi | \varphi \rangle = 1$. If each atom contributes one electron, the bond energy (BE) per electron is the difference between the molecular energy and the average of the atomic energies:

$$BE = E - \left(\frac{H_{11} + H_{22}}{2} \right) \rightarrow E = BE + \frac{H_{11} + H_{22}}{2}. \tag{4}$$

Substitution of Eq.(4) into Eq.(3) gives:

$$\left[BE + \frac{H_{11} + H_{22}}{2} \right] = c_1^2 H_{11} + c_2^2 H_{22} + 2c_1c_2H_{12} \tag{5}$$

The bond energy is then:

$$BE = c_1^2 H_{11} + c_2^2 H_{22} + 2c_1 c_2 H_{12} - \left[\frac{H_{11} + H_{22}}{2} (c_1^2 + c_2^2 + 2c_1 c_2 S_{12}) \right] \tag{6}$$

Substituting the Mulliken bond order, P_{12},[2,3] for $2c_1 c_2 S_{12}$ and consolidating terms gives:

$$BE = P_{12}\beta_{12} + (c_1^2 - c_2^2) \left(\frac{H_{11} - H_{22}}{2} \right), \tag{7}$$

where:

$$\beta_{12} = \left(\frac{H_{12}}{S_{12}} - \frac{H_{11} + H_{22}}{2} \right). \tag{8}$$

If we assume an Extended-Hückel-theory-like (EHT) relationship:[17]

$$H_{12}(R_{12}) \approx \lambda_{12} S_{12}(R_{12}), \tag{9}$$

with H_{12} proportional to S_{12}, then:

$$BE \approx P_{12} \left(\lambda_{12} - \frac{H_{11} + H_{22}}{2} \right) + (c_1^2 - c_2^2) \left(\frac{H_{11} - H_{22}}{2} \right), \tag{10}$$

where β_{12} is now independent of R_{12}. The first term in Eq.10 represents the covalent bond energy as a function of bond order (BEBO). The second term represents charge transfer between atoms 1 and 2 as expressed through the atomic populations. The Coulomb attraction between the charges induced by this charge transfer tends to cancel the charge transfer energy, except in the case of ion such as $Li^+ F^-$. If we set:

$$\lambda_{12} = 1.75 \frac{(H_{11} + H_{22})}{2}, \tag{11}$$

then Eq.11 represents EHT. The enormous (qualitative) success of EHT attests to the utility of this fundamental relationship between bond energy and bond order.

The BEBOP Model

We employ the linear relationship between the extended Hückel-type covalent bond energy and bond order:[17]

$$E_{\text{Bond}} = P_{12}\beta_{AB}, \tag{12}$$

where β_{AB} is the linear parameter for elements A and B, and P_{12} is the bond order between atoms 1 and 2.

The extended Hückel-type covalent bond energy method prominently appears in many SQM methods. SQM methods compute approximate electronic structure energies efficiently enough to be suitable for large-size systems (e.g., biochemistry, material science, macromolecules). The general formalism is similar to *ab initio* methods except that integrals are reduced drastically by implementing and parametrizing analytical expressions to fit accurate thermochemical data.[18] Many chemical applications[19–22] use SQM methods extensively for qualitative studies. The quantitative results of these methods, in general, are unreliable for compounds distinct from the training set from which parameters were obtained. This indicates that SQMs are not reliable for extrapolating to large-scale systems.

The BEBOP model attempts to extrapolate the total system atomization energies and bond energies using SQM-like formulas that employ reliable HF orbital populations. The current model includes hybridization and short-range repulsion to describe the energy of systems containing H, Li, Be, B, C, N, O and F. To generate the necessary population data, including P_{12}, ROHF/6-311+G(3d2f,2df,2p) calculations were used with the MinPop algorithm,[23] which projects an extended basis set onto a minimum basis set, providing a reliable Mulliken population analysis for bond orders and hybridization. However, the charges obtained are not always realistic. Since we observed that the charge transfer energy was generally balanced by the resulting electrostatic interaction energy, both terms were neglected. The parameters were determined by fitting restricted-open-shell (RO) model chemistry based on the complete basis set (CBS)-quadratic Becke3 (QB3), or ROCBS-QB3

for short, model of Wood and co-workers[24] to 36 reference species, so that we were not limited to experimentally known reference species. We were motivated to develop the BEBOP model as a bond energy analysis tool capable of emulating the accuracy of ROCBS-QB3 model for total atomization energies while also giving insights into other bond energy contributions. This method in its current form is only meant to be used for single point energy calculations on geometries using another computational method.

As the size of the molecules increased, inclusion of nonbonded steric effects became increasingly important. The steric interactions were described by the same linear relationship in Eq.12 with the negative bond orders. Unless otherwise stated, we elect to derive parameters for BEBOP for the geometry optimization obtained from UB3LYP/6-311G(2d,d,p) calculations. All ROCBS-QB3 and ROHF calculations were performed in GAUSSIAN16(Revision C.01).[25]

Orbital Hybridization

The concept of orbital hybridization was first introduced by Pauling.[1] He referred to the energy of hybridization as a change in quantization energy, which he calculated to be 9.3 eV for sp^3 hybridization of carbon. The BEBOP hybridization energy contribution was estimated from the $2s$ gross orbital occupations and the energies of appropriate atomic states. The energy due to hybridization of molecular carbon is taken as the number of $2s$ electrons in a ground state carbon atom minus the number of molecular $2s$ electrons, multiplied by the ^{3}P$\rightarrow$ ^{5}S atomic excitation energy (4.26 eV):

$$E(A)_{\text{hybridization}} = \left[n_{2\text{s}}(A)_{\text{ref}} - \sum_{i}^{\text{MO}} |C_{i2\text{s}_A}|^2 \right] \Delta E(A)_{2\text{s}\rightarrow 2\text{p}} \tag{13}$$

As a consequence of using the depletion of the $2s$ occupation to measure the $2s \rightarrow 2p$ hybridization, the additional energy required to remove an electron completely is the $2p$ ionization potential (IP).

Short-Range Repulsion

When the linear bond-energy/bond-order relationship is applied to a diatomic molecule, a fundamental deficiency becomes obvious. As nucleus 1 penetrates the exponentially decaying electron density around atom 2, the electron 2- nucleus 1 attraction is reduced and is no longer sufficient to cancel the nuclear-nuclear repulsion completely. Lennard-Jones[26] proposed that short-range repulsion could be expressed conveniently as a power law of the form $1/R^{12}$, but London showed that an exponential form is more appropriate.[27] Several exponential forms have been proposed.[28–33] A simple exponential form is employed in the BEBOP model:

$$E_{\text{repulsion}} = D_0 e^{-\zeta(R_e - R_\sigma)}. \tag{14}$$

This form ensures the correct energy at the classical turning point, R_σ, and the equilibrium geometry, R_e (Figure 1). The exponential parameters, ζ_{AB}, were evaluated by fitting the potential energy curves of the reference molecules, $H_nA{\sim}BH_m$ (*vide infra*). The ROHF BEBOP model describes the potential energy curves for pairs of monovalent atoms remarkably well, in spite of the improper dissociation of both the ROHF wave function and energy (Figure 1). The potential energy curves for polyvalent species such as carbon present a greater challenge. Rather than increase the complexity of the BEBOP model, we have elected to compromise and select a value for the exponent, ζ_{CC}, that reproduces the energies of ethane, ethylene, and acetylene, at their equilibrium geometries, R_e (Figure 2). Because the BEBOP model cannot explicitly calculate zero-point vibrational energies (ZPVE), we have employed the ROCBS-QB3 BDE value that includes ZPVE (i.e., D_0).

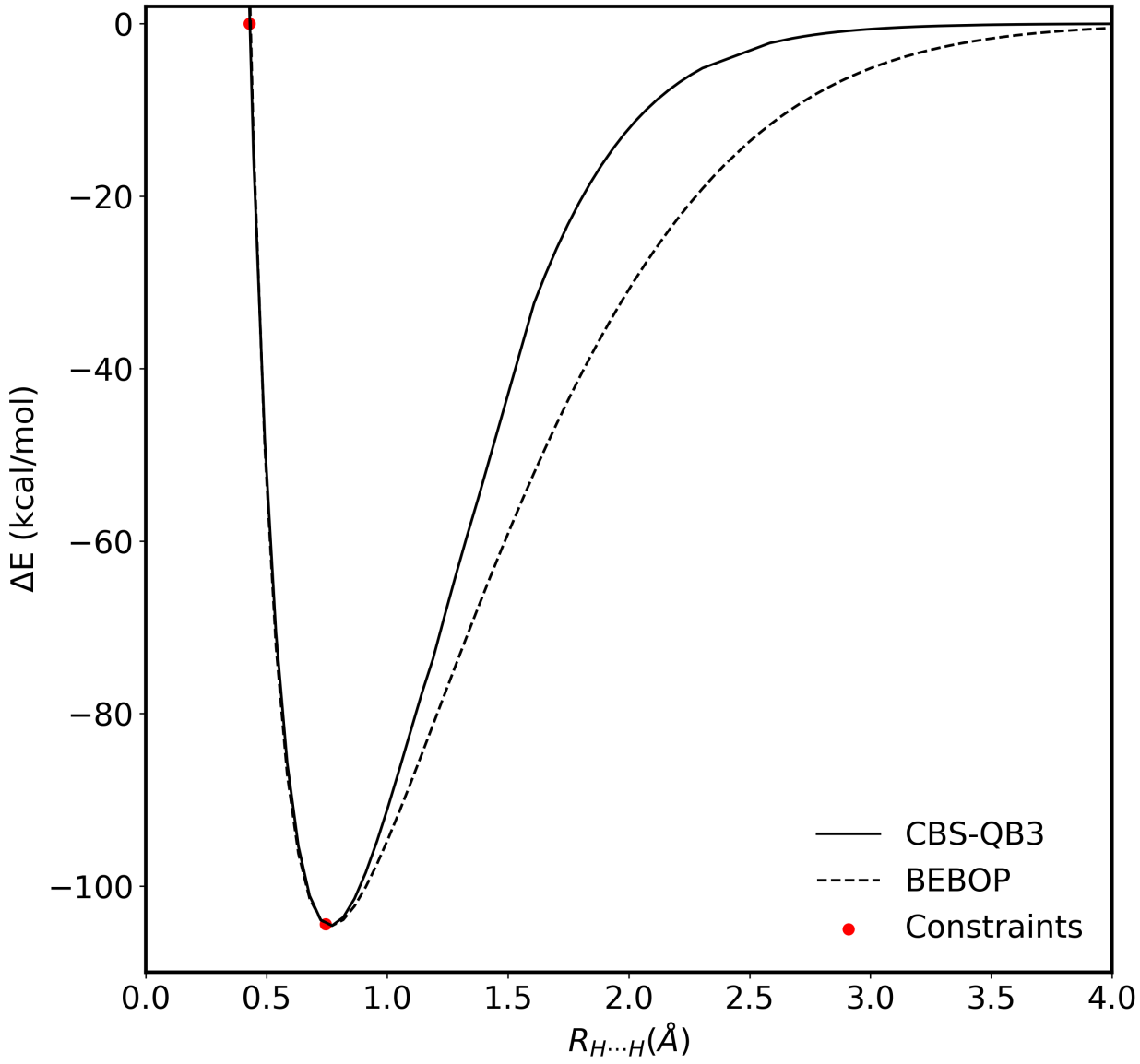


Figure 1: The BEBOP potential energy curve for H_2; $\beta = 144.77$ kcal/mol, $D_e = 104.45$ kcal/mol, $\zeta = 8.19$ Å^{-1},$R_\sigma = 0.462$ Å.

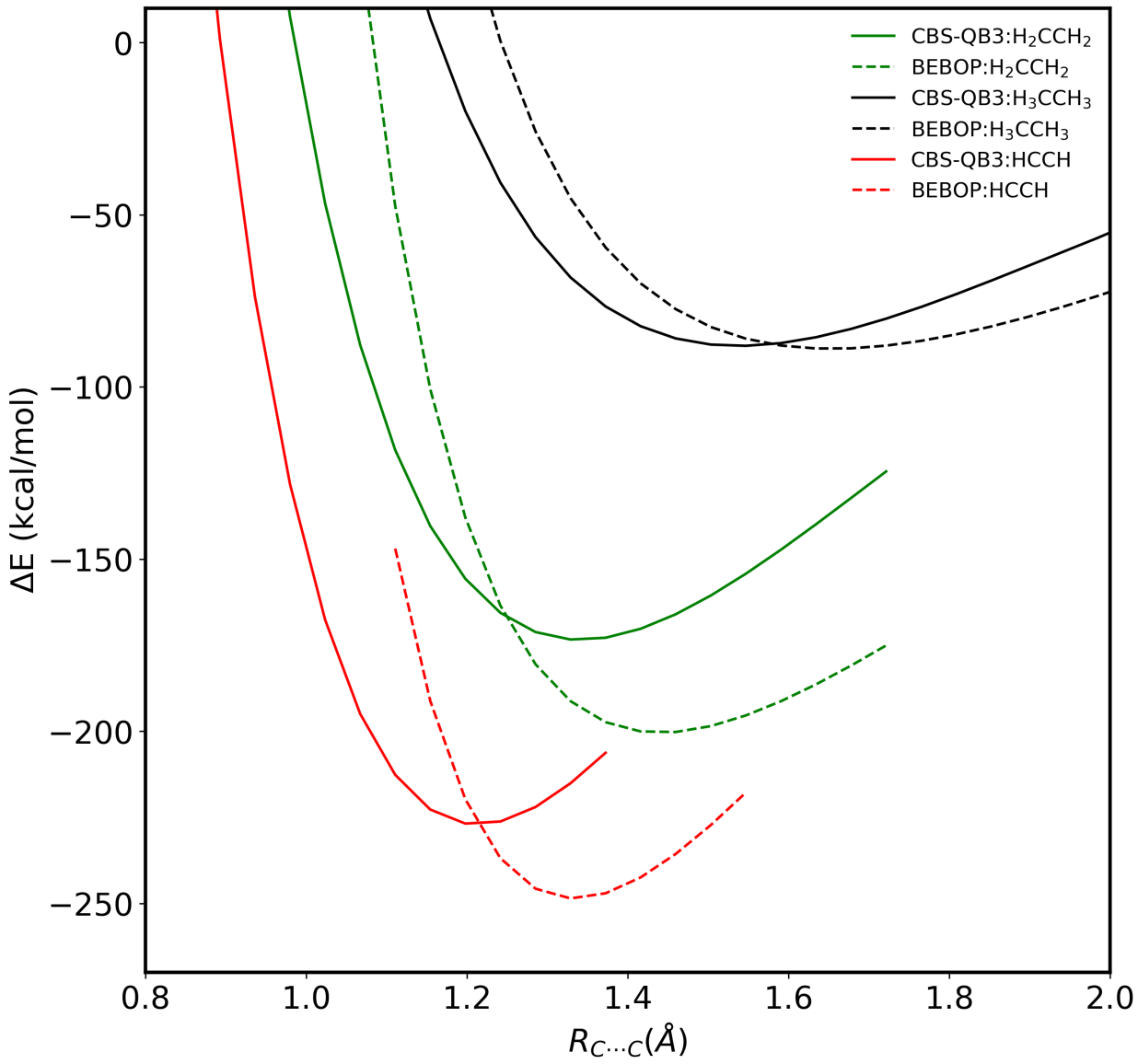


Figure 2: The BEBOP potential energy curves for ethane, ethylene, and acetylene.

The short-range repulsion term in Eq.14 plays a significant role in maintaining the viability of the linear BEBOP relationship expressed in Eq.12. Without this short-range repulsion, we could not describe the wide variety of C∼C bonds (Figure 3), unless we employed a totally unjustified nonlinear bond-energy/bond-order relationship. The physically necessary short-range-repulsion term is thus

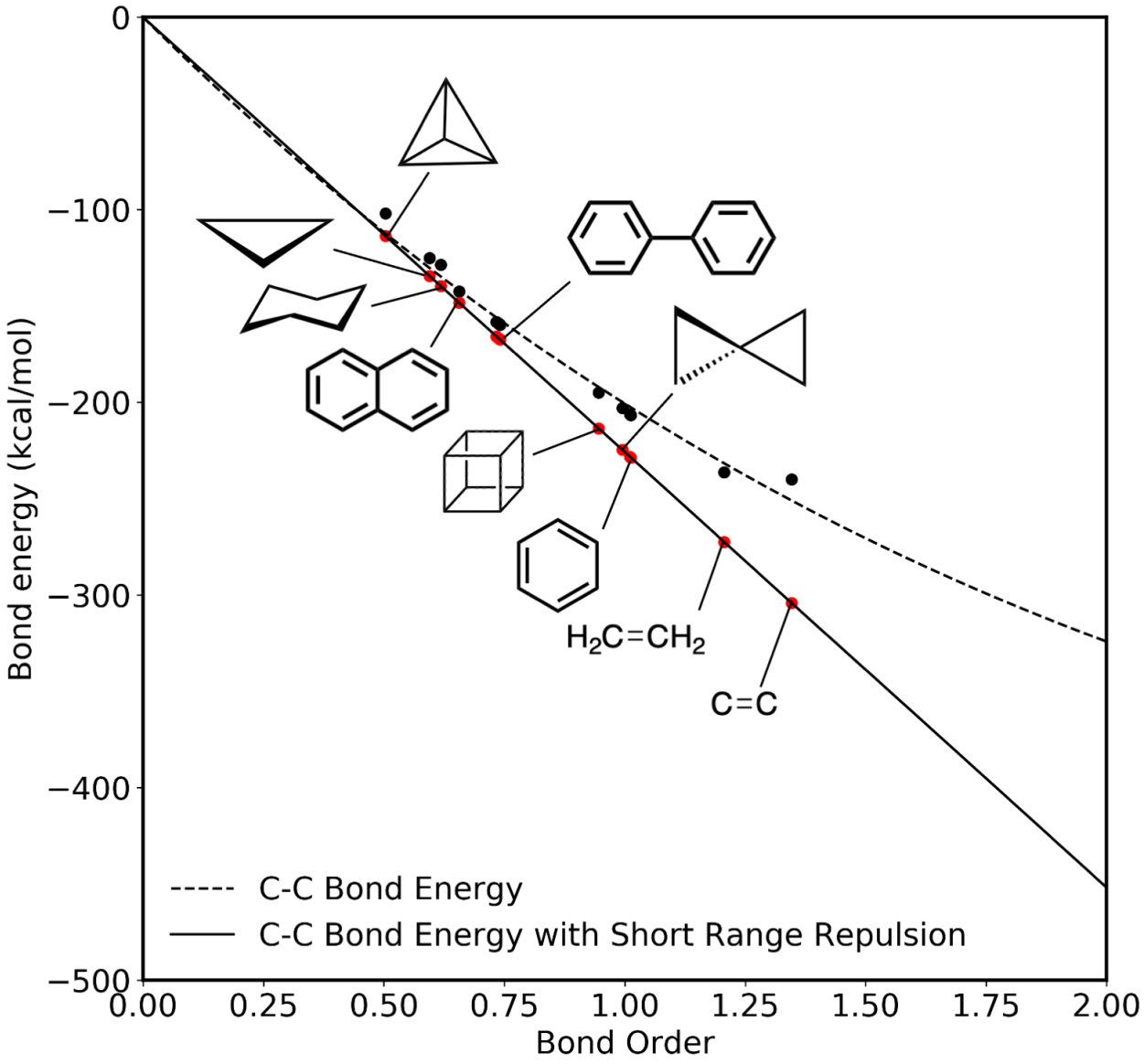


Figure 3: The relation between C∼C bond energies and bond orders for various hydrocarbons.

included for practical reasons, even if we do not attempt to explore potential energy surfaces. This is not the case for all atom pairs, AB. For example, we could describe C∼N bonds quite adequately without the short-range repulsion (Figure 4). The optimum value for ζ_{CN} is therefore very large ($\zeta_{\mathrm{CN}} = 13.84$ Å^{-1}), effectively removing any influence from the region around R_e.

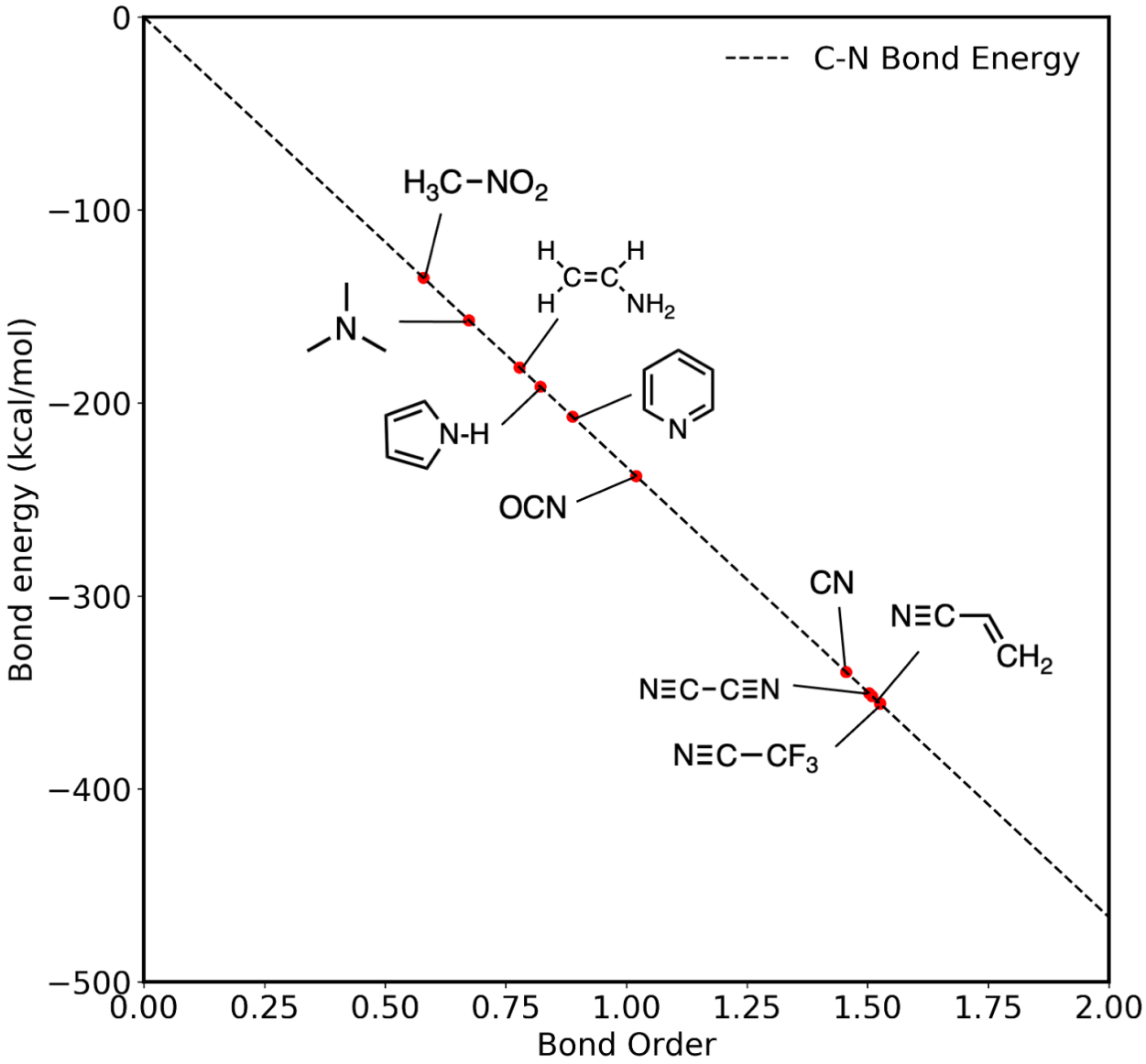


Figure 4: The relation between C∼N bond energies and bond orders across various hetero-organics.

The BEBOP model includes the attractive linear bond-energy/bond-order relationship in Eq.12, the short-range exponential repulsion in Eq.14, and the atomic hybridization energy in Eq.13. The model thus includes two fixed parameters (R_e[A,B] and D_0[A,B]), and two adjustable parameters (β_{AB} and ζ_{AB}) for each pair of elements, A and B.

BEBOP Parameters

The BEBOP parameters could be assigned using either experimental or ab initio benchmarks. We have elected to use total atomization energy from the ROCBS-QB3 model chemistry, $\Delta_{\rm AT}E_0$, with the minimum basis set Mulliken population analysis bond-orders and populations [GAUSSIAN16 command: # Pop=(MBS,Bonding)] derived from the largest SCF calculation employed in this model, ROHF/CBSB3 (CBSB3 is keyword in GAUSSIAN16 for 6-311+G(3d2f,2df,2p)). This is consistent with the intended role of BEBOP for the interpretation of *ab initio* energies. A choice of either experimental benchmarks or more accurate *ab initio* calculations (e.g. W*n* and G*n* models) would place greater restrictions on the range of accessible comparisons. A faster but less reliable benchmark could introduce spurious effects in the BEBOP bond energies.

The BEBOP model requires a hybridization parameter, $\Delta E(A)_{\rm 2s\rightarrow 2p}$, for each element, A (except hydrogen), and four parameters, β_{AB}, D_{AB}, $R_e(A,B)$ and ζ_{AB}, for each pair of elements, AB:

$$\Delta_{\rm AT}E_0 = \sum_{A\neq B}\left(\left[\sum_{\mu_A\nu_B}^{\rm AO}\sum_{i}^{\rm MO} C_{i\mu_A}C_{i\nu_B}S_{\mu_A\nu_B}\right]\beta_{AB} + D_{AB}\exp(-\zeta_{AB}\left[R_{AB}-\frac{R_e(A,B)}{\sqrt{2}}\right])\right) + \sum_{A}\left[n_{\rm 2s}(A)_{\rm ref} - \sum_{i}^{\rm MO}|C_{i{\rm 2s}_A}|^2\right]\Delta E(A)_{\rm 2s\rightarrow 2p} \tag{15}$$

The BEBOP hybridization parameters are assigned to be the ROCBS-QB3 atomic 2s→ 2p excitation energies given in Table 1. The reference molecules in Table 2 were used to determine the remaining parameters. In most cases, they represent the strongest bond that can be formed between the pair, A,B. However, in four cases (i.e., HBe–BeH vs. Be_4, HBe-OH vs. Be=O, HBe–F vs. FBe–F, and *cyclo*-BC_2H_3 vs. HB=CH_2), we found significant differences in parameter values for alternative reference choices and compromised with an average for the species in Table 2.

Table 1: ROCBS-QB3 ground and excited state energies.

Atoms	n_{2s}	n_{2p}	ROCBS-QB3[a]	$\Delta E_{2s\rightarrow 2p}$[b]	Source	
H			-0.4998179			
He			-2.9029321			
Li	1		-7.4320270	42.479	Li	$^2S \rightarrow {}^2P$
Be	2		-14.6205403	65.008	Be	$^1S \rightarrow {}^3P$
B	2	1	-24.6019582	85.177	B	$^2P \rightarrow {}^4P$
C	2	2	-37.7855234	98.098	C	$^3P \rightarrow {}^5S$
N	2	3	-54.520439	134.827	N^+	$^3P \rightarrow {}^5S$
O	2	4	-74.9879809	172.055	O^{2+}	$^3P \rightarrow {}^5S$
F	2	5	-99.643349	209.517	F^{3+}	$^3P \rightarrow {}^5S$

[a] in hartrees. [b] in kcal/mol.

Table 2: Reference species used to determine the ζ_{AB}, D_{AB}, $R_e(A,B)$, and D_{AB} parameters for BEBOP.

	H	Li	Be	B	C	N	O	F
H	H_2							
Li	LiH	Li_2						
Be	BeH_2	Li-BeH	HBe-BeH,Be_4					
B	BH_3	Li-BH_2	HBe-BH_2	cyclo-B_3H_3				
C	CH_4	Li-CH_3	HBe-CH_3	cyclo-BC_2H_3,HB=CH_2	H-C ≡ C-H			
N	NH_3	Li-NH_2	HBe-NH_2	cyclo-$B_3N_3H_6$	cyclo-C_5H_5N	HN=NH		
O	H_2O	Li−OH	HBeOH,Be=O	cyclo-$B_3O_3H_3$	H_2C=O	N=O	O=O=O	
F	HF	LiF	BeHF,BeF_2	HBF_2	CF_4	NF_3	HOF	F_2

The bond energy parameter, β_{AB} (Table 3), was constrained to reproduce the ROCBS-QB3 atomization energy, $\Delta_{\rm AT}E_0$, of the reference species. These parameters are in general qualitatively similar (i.e., ranging from 1/2 to 2 times) to the extended Hückel theory values, $(0.75[1/2(\mathrm{IP}_A + \mathrm{IP}_B)])$. However, the value of β_{AB} is artificially large for ion pairs because we have not accounted for the Coulomb attraction between the ions. Note that these bond energy parameters implicitly include ZPVE.

Table 3: BEBOP β parameters including ZPVE in kcal/mol.

	H	Li	Be	B	C	N	O	F
H	144.77							
Li	119.71	43.41						
Be	143.19	86.36	122.27					
B	168.44	121.82	147.88	160.70				
C	178.45	178.60	193.93	201.30	225.87			
N	192.39	207.00	217.45	233.62	233.16	215.55		
O	258.42	342.56	298.97	316.25	291.28	271.38	257.25	
F	372.61	760.99	474.24	468.04	403.33	322.98	252.10	289.78

The equilibrium bond length, $R_e(A,B)$ (Table 4), is the UB3LYP/6-311G(2d,d,p) optimized A,B distance (i.e., the geometry employed in the ROCBS-QB3 model) for the reference species. The best reference choice for $R_e(A,B)$ was generally the longest bond for the pair, A,B. This provided an adequate estimate of $R_\sigma \approx R_e/\sqrt{2}$ for 31 pairs, A,B, but the five exceptions indicated in Table 4 required adjustment to fit the short-range repulsion of the $H_nA{\sim}BH_m$ potential energy curve.

Table 4: BEBOP R_e parameters (Å).

	H	Li	Be	B	C	N	O	F
H	0.654[a]							
Li	1.593	2.705						
Be	1.327	2.399	2.080[a]					
B	1.190	2.183	1.867	1.727				
C	1.091	1.667[a]	1.673	1.553	1.532[b]			
N	1.016	1.569[a]	1.495	1.388	1.392[c]	1.095[d]		
O	0.962	1.570	1.518[a]	1.271	1.200[e]	1.148	1.207[f]	
F	0.920	1.560	1.373	1.324	1.389	1.430	1.434	1.408

[a] R_e adjusted to fit potential energy curve. [b] from H_3C-CH_3. [c] from H_2CCH-NH_2. [d] from $N \equiv N$. [e] from $H_2C = O$. [f] from $\cdot O{=}O\cdot$.

The BEBOP bond dissociation energy at 0 K parameter, D_0 (Table 5), is the ROCBS-QB3 $\Delta_{AT}E_0$ for the reaction: $H_nA{\sim}BH_m \rightarrow AH_n + BH_m$ of the reference species in Table 2.

Table 5: BEBOP bond dissociation energy with ZPE, D_0, parameters (kcal/mol).

	H	Li	Be	B	C	N	O	F
H	104.45							
Li	55.68	24.03						
Be	92.28	42.30	71.67					
B	104.54	44.48	82.44	136.32				
C	103.60	46.45	92.03	143.93	226.85			
N	105.93	72.02	120.86	177.14	157.24	122.65		
O	117.73	102.88	147.18	219.44	179.47	150.43	119.71	
F	136.01	136.27	176.64	169.55	110.06	69.31	48.44	37.44

In all cases of monovalent elements, the short-range repulsion exponent, ζ_{AB} (Table 6), was adjusted to fit the potential energy curve as in Figure 1. Elements capable of multiple bonds required a compromise as shown for C∼C bonds in Figure 2. We could fit several species at their equilibrium bond length reasonably well, but the potential energy curves were only qualitatively correct. The BEBOP parameters in Tables 1, 3, 4, 5, and 6, were used for all calibration and application studies.

Table 6: BEBOP ζ_{AB} parameters (Å^{-1}).

	H	Li	Be	B	C	N	O	F
H	8.19							
Li	2.71	3.26						
Be	5.31	2.85	4.28					
B	5.92	5.77	4.27	6.73				
C	7.17	7.02	4.26	6.68	7.53			
N	7.81	4.85	4.25	6.64	7.44	13.84		
O	8.68	2.18	4.25	5.57	7.30	8.99	8.42	
F	9.56	1.24	4.99	4.61	7.57	7.50	10.91	1.17

Bond Energies vs. Bond Dissociation Energies

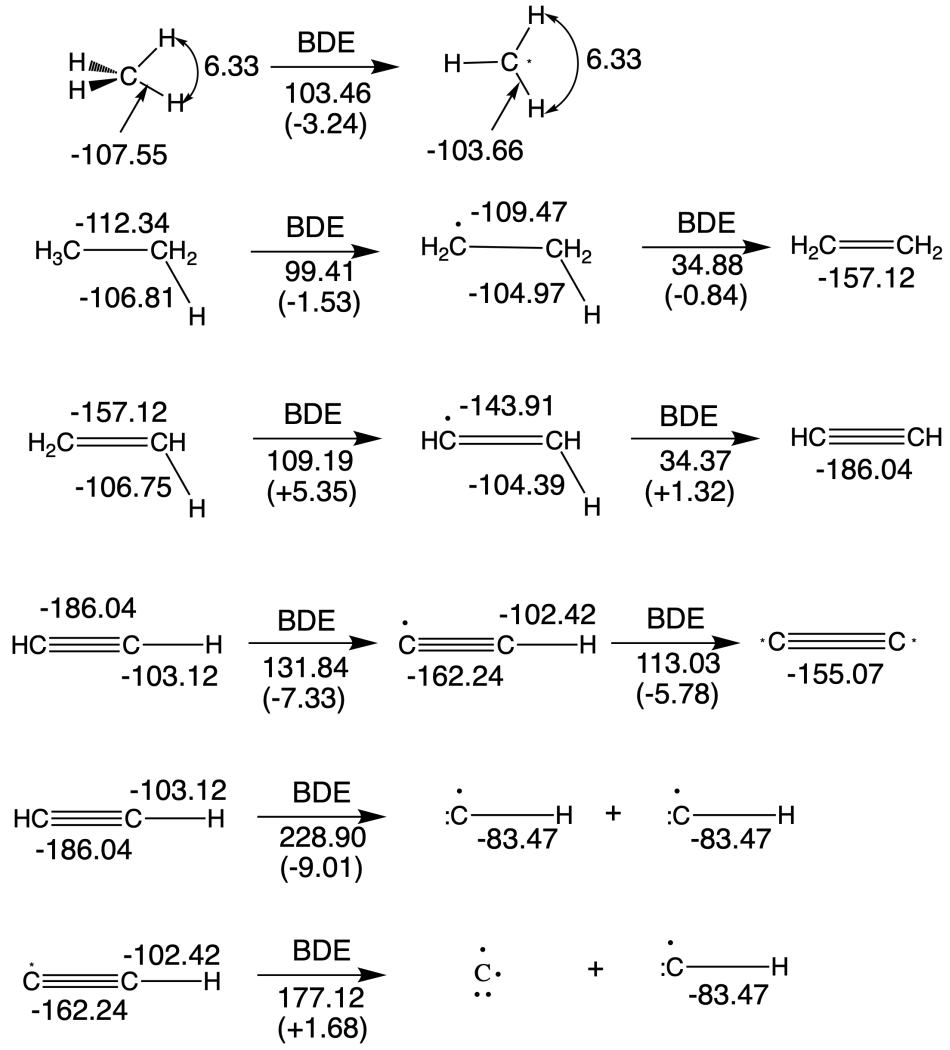


Figure 5: Deviation (Theory-Experiment) of BEBOP's bond dissociation energy (shown in parentheses) to the experimental bond dissociation energies from an active thermochemical table (ATcT[34,35]) found in Ref. 36 for various small hydrocarbons. The BEBOP net bond energies for each chemical species are shown next to the chemical bonds, and curly arrows represent anti-bonding interactions. All values are shown in kcal/mol.

Bond dissociation energies (or enthalpies) are simply the energy necessary to break a bond. Bond energies can be defined as the contribution a bond makes to the total energy of a molecule relative to the energy of the separated atoms. The two are necessarily equal for diatomic molecules, but are not in general equal for polyatomic molecules. For example, the energy necessary to separate a methane molecule into four hydrogen atoms and a carbon

atom is 392.3 ± 0.2 kcal/mol. The bond energy for each of the four C–H bonds is therefore -98.1 kcal/mol. However, the $CH_4 \rightarrow CH_3 + H$ bond dissociation energy is 103.35 kcal/mol (which includes the geometric relaxation energy of the methyl radical). The C-H bond energy has decreased to -96.20 kcal/mol in CH_3. The bond dissociation energy (BDE), in general, can be interpreted as the difference between the sum of the relaxed product bond energies and the sum of the reactant bond energies. The BEBOP model includes the repulsive geminal $H \cdots H$ interactions arising from the typically negative $1,3$ bond orders. The BEBOP value for the $CH_4 \rightarrow CH_3 + H$ bond dissociation energy is therefore $[3 \times 6.33 + 3 \times -103.66] - [6 \times 6.33 + 4 \times -107.55] = 100.23$ kcal/mol (Figure 5). Note that our convention results in negative bond energies, but positive bond dissociation energies.

The sequence of C-H bond dissociation energies from ethane to C_2 provides a more interesting bond energy analysis (Figure 5). The C–H bond dissociation energies vary by almost 100 kcal/mol (from 34.25 kcal/mol to 131.44 kcal/mol), but the C–H bond energies are almost constant (from -102.42 kcal/mol to -106.81 kcal/mol). The principal variation in bond orders and thus bond energies resides in the C∼C bonds. The large increase in C∼C bond energy (and resulting small C–H BDE) when a new π-bond forms is hardly a revelation, but it is not generally appreciated that the large C–H BDE of acetylene is a consequence of the much weaker C≡C bond in ethynyl radical, rather than an especially strong C–H bond in acetylene.

The C≡C bond dissociation energies from acetylene and the ethynyl radical (Figure 5) demonstrate that the reduced strength of the C≡C bond is not an artifact of the BEBOP analysis. These C≡C bond dissociation energies are enhanced by the small bond energy of the C–H product, but the reduced C≡C bond energy of the ethynyl radical is unmistakable.

1,3-Antibonding Repulsions

The ubiquitous appearance of repulsive geminal interactions in the BEBOP analysis invites some explanation. The transition state for the reaction of H_2 with a hydrogen atom provides

a convenient prototype. Since the singly occupied $1\sigma_u$ orbital must be orthogonal to the $1\sigma_g$ orbital, the $1\sigma_u$ orbital will necessarily introduce a 1, 3-antibonding interaction. This repulsive germinal interaction is responsible for the linearity of the H_3 transition state, and is a recurrent theme in hundreds of molecules we have examined. These 1, 3-antibonding interactions are generally $\sim$ 6 kcal/mol for H$\cdots$H interactions, $\sim$ 7 kcal/mol for C$\cdots$H interactions, and $\sim$ 8 kcal/mol for C $\cdots$ C interactions. They are reduced to $\sim$ 3 kcal/mol for linear structures, but are amplified by about a factor of 2 to 3 in transition states such as the H_3 structure below (Figure 16).

The BEBOP Analysis

A detailed analysis of bond energies must include the hybridization energy:

$$\begin{aligned}\Delta_{\mathrm{AT}}E_0 &= \sum_{A\neq B}\left[\left(\sum_{\mu_A\nu_B}^{\mathrm{AO}}\sum_{i}^{\mathrm{MO}} C_{i\mu_A}C_{i\nu_B}S_{\mu_A\nu_B}\right)\beta_{A,B} + D_{AB}\exp(-\zeta_{AB}\left[R_{AB}-\frac{R_e(A,B)}{\sqrt{2}}\right])\right] \\ &+ \sum_{A}\left[n_{2\mathrm{s}}(A)_{\mathrm{ref}} - \sum_{i}^{\mathrm{MO}}|C_{i2\mathrm{s}_A}|^2\right]\Delta E(A)_{2\mathrm{s}\rightarrow 2\mathrm{p}} \\ &= \sum_{A\neq B}E(A,B)_{\mathrm{gross}} + \sum_{A}E(A)_{\mathrm{hybridization}}\end{aligned} \tag{16}$$

It is often desirable to partition the total energy over the bonds, without terms associated with individual atoms. We have elected to partition the hybridization energy for an atom

over all bonds to this atom, weighted by the *gross* bond energies:

$$E(A,B)_{\text{net}} \equiv E(A,B)_{\text{gross}} + E(A)_{\text{hybridization}} \left[\frac{E(A,B)_{\text{gross}}}{\sum_{C \neq A} E(A,C)_{\text{gross}}} \right] + E(B)_{\text{hybridization}} \left[\frac{E(A,B)_{\text{gross}}}{\sum_{C \neq B} E(B,C)_{\text{gross}}} \right] \tag{17}$$

$$\Delta_{\text{AT}} E_0 = \sum_{B \neq A} E(A,B)_{\text{net}}$$

These molecular energy components can be concisely presented as a square matrix as shown in Figure 6. We shall employ the convention that the diagonal elements are the hybridization energies, the elements above the diagonal are the *gross* bond energies, and the elements below the diagonal are the *net* bond energies.

Table 7: Slater atomic exponential exponents ζ_{Slater} for valence orbitals fitted to STO-6G minimal basis.

Atoms	Valence Orbitals	ζ_{Slater} (Å^{-1})
H	1s	1.23
He	1s	1.67
Li	2s	0.80
Be	2s	1.15
B	2p	1.50
C	2p	1.72
N	2p	1.95
O	2p	2.25
F	2p	2.55

The BEBOP model can seamlessly decompose the individual bond energies into σ and π bond contributions. The model separates the total bond order to σ and π bond orders by employing: (i) Roothan's two-center σ-σ and π-π overlap integrals[37] and (ii) fixed Slater atomic exponents (ζ_{Slater}) for valence electrons. The ζ_{Slater} is an essential parameter to compute the overlap integrals. We generated the optimum values (Table 7) by fitting a

Slater function to STO-6G atomic basis sets for each first row elements' valence orbitals.[38,39] The gross and net bond energies for individual σ and π bonds are then computed using Eq.16 and Eq.17, respectively. Eventually, the BEBOP model generates the individual σ and π molecular bond energy components as elements from lower-diagonal matrices.

C-H $^2\Pi$

	C	H
C	14.19	-97.66
H	-83.47	0.00

C-H $^4\Sigma$

	C	H
C	60.88	-114.14
H	-53.26	0.00

H−C≡C·

	H	C	C
H	0.00	-128.60	3.18
C	-102.42	88.00	-303.62
C	2.35	-162.24	78.73

H−C≡C−H

	H	C	C	H
H	0.00	-129.10	3.47	-0.10
C	-103.12	87.96	-311.39	3.47
C	2.77	-186.04	87.96	-129.10
H	-0.10	2.77	-103.12	0.00

Figure 6: BEBOP's bond energy analysis matrices depicting the gross bond energies (upper-diagonal elements), hybridization energies (diagonal elements), and net bond energies (lower-diagonal elements) values for methylidyne radicals (C−H $^2\Pi$ and $^4\Sigma$ state), ethynyl radical (H−C≡C:), and acetylene (H−C≡C−H).

The C–H diatomic molecule nicely illustrates the significance of each energy component (Figure 6). The $^2\Pi$ ground state bond is essentially a combination of a carbon 2p with a hydrogen 1s, so the hybridization energy is very small (14.19 kcal/mol), but the bond is relatively weak ($E_{\text{gross}} = -97.66$ kcal/mol). The carbon of the $^4\Sigma$ excited state bonds with an *sp* hybrid orbital, forming a stronger bond ($E_{\text{gross}} = -114.14$ kcal/mol), but the enhanced bonding is insufficient to overcome the increased hybridization energy (60.89 kcal/mol), so the $^2\Pi$ state net bond energy (−83.6 kcal/mol) is greater than the $^4\Sigma$ state net bond energy (−53.26 kcal/mol). The gross bond energy is a more direct measure of the orbital bonding, but the net bond energy is a more complete measure of the bonding energy. The BEBOP bond energy for the $^2\Pi$ state is in good agreement with the ROCBS-QB3 result (−80.08

kcal/mol), but without a correction for the reduced repulsion between electrons with parallel spin, BEBOP underestimates the $^4\Sigma$ bond energy (ROCBSQB3: $D_0 = 61.31$ kcal/mol).

We now have the tools to analyze the reduced C≡C bond energy of the ethynyl radical relative to acetylene (Figure 6). The gross C≡C bond energy of the ethynyl radical (−303.62 kcal/mol) is just slightly smaller (than −311.39 kcal/mol in acetylene), consistent with the very small increase in C≡C bond length from 1.198 Å to 1.202 Å. The hybridization energy of the terminal carbon (78.73 kcal/mol) is more than 9 kcal/mol smaller for the ethynyl radical, but the terminal carbon lacks a C–H bond to share this hybridization energy. The net C≡C bond energy of the ethynyl radical (−162.24 kcal/mol) is therefore 23.80 kcal/mol smaller than that of acetylene (−186.04 kcal/mol).

We can rescale the BEBOP net bond energies to reproduce the ROCBS-QB3 energies exactly (Figure 7), but the agreement is sufficiently good before rescaling that our interpretation remains unchanged.

	H	C	C
H	0.00		
C	-99.75	0.00	
C	2.29	-158.01	0.00

	H	C	C	H
H	0.00			
C	-103.12	0.00		
C	2.77	-186.13	0.00	
H	-0.10	2.77	-103.16	0.00

Figure 7: Rescaled BEBOP net bond energy matrices to ROCBS-QB3 energies for ethynyl radical and acetylene molecules.

The hybridization energy is generally ∼ 75 kcal/mol for sp^3 hybridized carbons, ∼ 82 kcal/mol for sp^2 hybridized carbons, and ∼ 90 kcal/mol for sp hybridized carbons, increasing as more pure p orbitals are occupied.

Bond energies and charges of atoms in molecules are useful intuitive concepts that have enjoyed wide acceptance as qualitative parameters, but cannot be given unique quantitative definitions. Just as the only rigorous constraint on "atomic charges" is that the sum of these charges equal the total charge of the molecule, the only rigorous constraint on bond energies is that the sum of the bond energies equal the total energy required to dissociate the molecule

into the constituent atoms (i.e., the atomization energy).

Computational Details

We have employed a calibration set of 192 species including the subset (109 species) of the G2/97 test set[40] containing only the elements H through F (Table 8). The errors in the total atomization energies at 0 K, $\Delta_{\mathrm{AT}}E_0$, were determined by comparison with available experimental data or by comparison with ROCBS-QB3 energies in the absence of reliable experiments. The experimental $\Delta_{\mathrm{AT}}E_0$ used for our reference were at room temperature and were converted to 0 K using the thermal corrections from ROCBS-QB3's thermochemistry calculations.

We will now give the computational details of our atomization energies at 0 K and BDE calculations. We used various computational methods to compare BEBOP's performance for energetics across the G2/97 test set. In the Results and Discussion sections, we will use the following abbreviations to compare the performance of BEBOP to other computational methods: root-mean-squared error (RMSE), mean absolute deviation (MAD), and maximum unsigned deviation (MAX).

Calibration for Atomization Energies

We compute the atomization energy at 0 K as the difference in the sum of the molecular total electronic energy,$E_{\mathrm{mol.,elec}}$, and zero-point vibrational energy,$E_{\mathrm{mol.,ZPVE}}$, to the sum of the constituent atomic energies,$E(i)$, present in the molecule:

$$\Delta_{\mathrm{AT}}E_0 = E_{\mathrm{mol.,elec}} + E_{\mathrm{mol.,ZPVE}} - \sum_i E(i). \tag{18}$$

Unless stated otherwise, the $E(i)$ and $E_{\mathrm{mol.,elec}}$ were computed for molecules and atoms, respectively, at the employed level of theory, but the $E_{\mathrm{mol.,ZPVE}}$ are from ROCBS-QB3's

thermochemistry calculations.

The ROCBS-QB3 method was computed in GAUSSIAN16[25] (i.e, using "# ROCBSQB3" as the keywords). This model attempts to estimate the total CCSD(T)/CBS energy using multiple quantum chemistry models. The composite method includes the following computations: (i) UB3LYP/6-311G(2d,d,p) for geometry optimization and frequency calculations (ZPVEs are scaled by 0.9900), (ii) UMP2/6-311+G(3d2f,2df,2p) for energy and CBS extrapolation calculations, (iii) UMP4(SDQ)/6-31+G(d(f),p) for energy calculations, and (iv) UCCSD(T)/6-31+G† for energy calculations. The calculations from steps (ii)-(iv) were done at the optimized geometry from (i), and no imaginary modes were present in any of our data set.

The BEBOP method required a geometry optimization calculation from a different QC method. We will use UB3LYP/6-311G(2d,d,p) as the recommended method to optimize our calibration data sets since the parameters had been generated at that method's geometry. We then run ROHF/6-311+G(3d2f,2df,2p) and the Mulliken minimum population analysis, so the minimal population Mulliken Bond orders and the 2s orbital populations are calculated. Finally, we compute the BEBOP atomization energy using Eq.15 with the parameters in Table 1 through Table 6.

We compare our model's performance on accuracy to various KS-DFT hybrids. We use B3LYP,[41] PBE0,[42–44] APF[45] and M062X[46] exchange-correlations. The geometry of the calibration set was done at the respective level of the method using Pople's 6-31G*[47–51] basis set. Single-point energies were then computed using Dunning's cc-pVTZ[52,53] basis sets on the optimized geometries, except PBE0/6-31G* geometries. We included dispersion corrections to each hybrid model (except APF and M062X) using Grimme's D3[54] empirical dispersion model with Becke-Johnson damping (D3BJ[55]). For the APF model, we used the empirical Petersson-Frisch dispersion model, which is designated as APFD. All DFT calculations were done in the GAUSSIAN16(Revision C.01) code.[25]

We include in our study the AM1,[56] PM6,[57] PM7,[58] and DFTB3[59,60] energies to provide

context for the accuracy of the BEBOP model by comparison with popular SQM models. All geometry optimization and single-point energy calculations were done at the respective SQM method in MOPAC16[61] (for AM1, PM6, and PM7) and DFTB+[62] (DFTB3 with the recommended 3OB[63,64] Slater-Koster element-pair parameters) codes. The AM1, PM6, and PM7 total energies from MOPAC16 include ZPVE corrections for molecules. For DFTB3,however, $E(i)$ were calculated using the recommended atomic energies found in the supporting information of Ref. 64 and the ROCBS-QB3's ZPVE values was used as $E_{\mathrm{mol.,ZPVE}}$.

The AM1 and DFTB3 model has a smaller number of data sets used. The AM1 model present in MOPAC16 does not have parameters for B, and DFTB3 does not have parameters for Li,Be, and B. Moreover, the DFTB3-3OB model's parameters are optimized for ground-state energies of closed shell species, so we discarded any excited-state species from our test sets. These excited-state compounds include C−H ($^4\Sigma$), CH_2 ($^3B_1$), C−F ($^4\Sigma$), CF_2 ($^3B_1$), and Be−H ($^2\Sigma$).

Bond Dissociation Energies

We made a data set of 79 BDEs for this investigation. Of the total BDEs, 30 BDEs were computed from the calibration set consisting of diverse chemical bonding(Figure 10). We elected to compute these diverse BDEs using the same computational approach and level of theory methods used in our atomization calibration studies. For a general homolytic bond dissociation reaction $H_nA{\sim}BH_m \rightarrow AH_n + BH_m$, BDEs at 0 K,$D_0$, are computed as:

$$D_0 = \Delta_{\mathrm{AT}}E_0(\mathrm{AH}_n) + \Delta_{\mathrm{AT}}E_0(\mathrm{BH}_m) - \Delta_{\mathrm{AT}}E_0(\mathrm{H}_n\mathrm{A} \sim \mathrm{BH}_m), \tag{19}$$

where $\Delta_{\mathrm{AT}}E_0(\mathrm{AH}_n)$ and $\Delta_{\mathrm{AT}}E_0(\mathrm{BH}_m)$ are the atomization energies (Eq.18) of the *relaxed* (i.e., ground-state geometry and energy) product states and $\Delta_{\mathrm{AT}}E_0(\mathrm{H}_n\mathrm{A} \sim \mathrm{BH}_m)$ is the atomization energy of $H_nA{\sim}BH_m$.

The other 49 BDEs are from small alcohol molecules (Figure 11) not present in our

calibration data set. These BDEs at room temperature (D_{298}) were computed for our familiar reaction scheme by summing D_0 and the thermal correction difference in the relaxed product states ($E_{\text{thermal}}(\text{AH}_n)$ and $E_{\text{thermal}}(\text{BH}_m)$) and the reactants($E_{\text{thermal}}(\text{H}_n\text{A} \sim \text{BH}_m)$):

$$D_{298} = D_0 + E_{\text{thermal}}(\text{AH}_n) + E_{\text{thermal}}(\text{BH}_m) - E_{\text{thermal}}(\text{H}_n\text{A} \sim \text{BH}_m). \tag{20}$$

The thermal corrections from this investigation were obtained from ROCBS-QB3 calculations. The computational methods for obtaining atomization energies at 0 K for the products and reactant molecules were the same as the atomization calibration study, and no imaginary frequencies from ROCBS-QB3 calculations were found.

We computed D_{298} data from a popular and free online machine learning model called the machine-learning derived, fast, accurate bond enthalpy tool (ALFABET).[65,66] The ALFABET model can predict and extrapolate D_{298} energies based on reference D_{298} data of small organic molecules at the M062X level of theory with Weigend's def2-TZVP[67] basis sets. The model can only compute D_{298} data for species consisting of C, H, O, and N atoms. By inserting a SMILES string or using a drawing tool in `https://bde.ml.nrel.gov`, ALFABET generates D_{298} values (from least to greatest) at low cost. The errors of this model relative to experiments are low for most compounds, which makes this model useful for gas-phase studies and is competitive to other computational models.

All D_{298} calculations were compared to Oyeyemi et al. multi-reference averaged coupled-pair functional theory (MRACPF2[68,69]) study on small alcohol and aldehyde molecules' D_{298} trends.[70] The errors were studied for C–C, C–H, C–O, and O–H alcohol bonds for all computational methods to gauge BEBOP's estimations for D_{298}. The number of sets present for each bond are 13, 20, 8, and 8, respectively.

Results and Discussion

As shown in Table 8, BEBOP outperforms all SQMs and several hybrid DFT methods for atomization energies. The RMSE (8.42 kcal/mol) is intermediate to M062X/6-31G* (7.11 kcal/mol) and APF/cc-pVTZ (9.82 kcal/mol). The MAD value (5.87 kcal/mol) is intermediate to APF/cc-pVTZ (5.62 kcal/mol) and APFD/cc-pVTZ (6.40 kcal/mol). BEBOP and other methods' performance on hydrocarbons, hetero-carbons, and inorganics will be discussed below. We will then discuss the results for the bond dissociation energies at 0 K and 298 K for all methods used in this study.

Hydrocarbons

The C∼C bonds of the hydrocarbons provide the broadest range of bond orders and bond energies and thus provide the most rigorous test of the BEBOP concept. We supplemented the 41 hydrocarbons and radicals of the G2/97 test set with 15 additional species such as the C−H $^4\Sigma$ excited state (to test the BEBOP treatment of hybridization), calicene (to test the ability of the BEBOP model to handle charges), and tetrahedrane (to push the limits of ring strain).

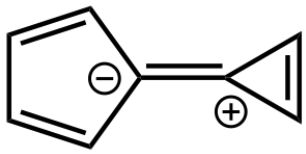
The BEBOP model gives similar or lower RMSE and MAD errors for hydrocarbons compared to hybrid DFT methods that uses large basis sets (see Table 8). On one hand, the BEBOP RMSE (4.80 kcal/mol) is intermediate between the ROCBS-QB3 RMSE value(2.37 kcal/mol) and the M062X/cc-pVTZ//M062X//6-31G* RMSE value(5.03 kcal/mol). On the other hand, BEBOP's MAD value (3.82 kcal/mol) is intermediate between B3LYP-D3BJ/cc-pVTZ//B3LYP-D3BJ/6-31G* value (2.82 kcal/mol) and M062X/6-31G*//M062x/6-31G* (3.86 kcal/mol). BEBOP is far more accurate than any other popular SQM method for hydrocarbon species. This is essential since the BEBOP formulation is similar to SQM formalism, except that the BEBOP model computes highly robust bond orders by using a minimal Mulliken population analysis on ROHF orbital populations.

Table 8: Root-mean-squared error (RMSE), maximum unsigned deviation (MAX), and mean absolute deviation (MAD) of the total atomization energies, $\Delta_{\mathrm{AT}}E_0$ (in kcal/mol) for the G2/97 test set and reference species.

Method	RMSE				MAX				MAD			
	Overall	hydrocarbons	heterocarbons	inorganics	Overall	hydrocarbons	heterocarbons	inorganics	Overall	hydrocarbons	heterocarbons	inorganics
composite thermochemical												
CBS-QB3//B3LYP/CBSB7	1.76	2.37	1.17	1.25	4.66	4.66	3.01	3.27	1.29	1.04	0.96	0.88
BEBOP//B3LYP/CBSB7	8.42	4.80	10.40	8.77	33.88	13.57	25.19	33.88	5.87	3.82	8.49	5.42
SQM												
DFTB3/3OB//DFTB3/3OB	14.85	8.45	10.01	29.01	74.97	25.64	36.08	74.97	11.00	5.61	6.65	20.66
PM7//PM7	17.13	15.00	10.96	21.98	84.72	39.27	24.15	84.72	9.11	5.25	5.87	13.31
AM1//AM1	20.64	14.50	15.70	30.20	170.33	36.42	48.92	170.33	10.16	9.13	8.58	13.85
PM6//PM6	27.24	17.86	11.48	39.32	275.50	59.10	32.48	275.50	12.11	6.39	6.50	19.54
DFT												
M062X/cc-pVTZ//M062X/6-31G*	5.36	5.03	2.61	7.00	30.88	26.74	6.70	30.88	3.54	2.55	1.99	4.90
B3LYP-D3BJ/cc-pVTZ//B3LYP/6-31G*	5.75	5.22	3.33	7.42	29.39	23.88	10.54	29.39	3.32	2.82	2.33	4.34
B3LYP-D3BJ/6-31G*//B3LYP/6-31G*	6.94	6.99	3.99	8.57	31.5	24.23	10.06	31.5	4.73	4.34	3.00	6.15
B3LYP/6-31G*//B3LYP/6-31G*	6.95	7.91	4.57	7.72	29.72	25.02	14.56	29.72	4.84	3.94	3.00	5.69
B3LYP/cc-pVTZ//B3LYP/6-31G*	7.02	10.16	3.65	6.14	27.61	25.68	12.01	27.61	4.71	5.99	2.29	3.91
M062X/6-31G*//M062X/6-31G*	7.11	5.76	4.84	9.22	33.6	25.91	13.16	33.6	5.12	3.86	3.74	7.13
APF/cc-pVTZ//APF/6-31G*	9.82	12.97	4.32	10.26	66.64	34.67	13.27	66.64	5.62	8.14	2.90	5.55
PBE0/6-31G*//PBE0/6-31G*	9.94	13.40	5.75	9.55	35.05	35.05	14.01	32.51	6.78	8.71	4.51	7.07
APF/6-31G*//APF/6-31G*	11.42	14.56	5.52	12.24	70.16	39.06	13.69	70.16	7.16	8.97	4.23	7.79
APFD/cc-pVTZ//APFD/6-31G*	11.57	16.08	5.02	11.31	66.8	43.93	15.11	66.8	6.40	9.41	3.04	6.00
APFD/6-31G*//APFD/6-31G*	13.19	17.97	6.20	13.08	70.33	48.38	15.7	70.33	8.23	10.79	4.60	8.10
PBE0-D3BJ/6-31G*//PBE0/6-31G*	15.43	24.80	7.62	10.04	69.2	69.2	22.3	34.29	9.61	14.42	5.12	7.52
B3LYP/CBSB7//B3LYP/CBSB7	28.76	14.55	6.98	43.69	269.41	38.11	19.22	269.41	12.41	7.59	2.68	25.19

Thus, our BEBOP model produces smaller errors for hydrocarbons species with negligible dispersion than the other methods computed in this paper. The BEBOP absolute error is intermediate with the MAX error of ROCBS-QB3 (4.66 kcal/mol) and B3LYP-D3BJ/6-31G*//B3LYP/6-31G* (23.88 kcal/mol). The MAX error for BEBOP's hydrocarbon data (13.57 kcal/mol) comes from the C_2 diradical. Similarly, this species gives MAX errors for B3LYP(all in Table 8, except B3LYP/cc-pVTZ energies), DFTB3, and M062X (for all shown in Table 8). The PM6 and PM7 MAX errors stem from cubane, while the AM1 MAX error is from *n*-octane. Finally, the PBE0-D3BJ/6-31G* and PBE0/6-31G* energies produce the MAX errors from anthracene and phenanthrene, respectively. The C_2 diradical has multireference behavior, which means that single determinant methods are not able to describe the physics of this molecule.

The moderate BEBOP error for calicene (7.79 kcal/mol):



is particularly noteworthy, since we omitted an explicit treatment of charge transfer and electrostatic effects. The large dipole moment (5.1 Debye) of calicene implies an essentially zwitterionic structure.

Hetero-organic Species

The BEBOP RMSE and MAD errors for hetero-organic species are higher, as the current BEBOP model give a poor description of polar bonds. DFT methods, however, outperform the current BEBOP model for hetero-organic molecules, as the MAD and RMSE errors are lower for the latter. The RMSE value for BEBOP (10.40 kcal/mol) is intermediate between DFTB3 (10.01 kcal/mol) and PM7 (10.96 kcal/mol). However, the MAD value (8.49 kcal/mol) is intermediate between DFTB3 (6.65 kcal/mol) and AM1 (8.58 kcal/mol) MAD

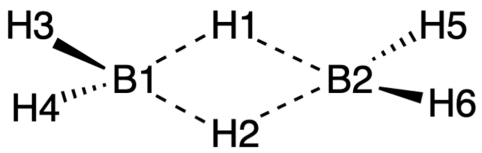
values. Thus, the BEBOP's performance for hetero-organics is similar to the performance of semi-empirical methods.

The BEBOP MAX value for hetero-organic species is attributed to weak description of C=O· molecule, and the BEBOP and DFT models significantly overestimates the energy of =C−F bonds, but has no difficulty with the fragments of CF_4 (CF, CF_2, etc.) as AM1, PM6, and PM7 methods. The AM1, PM6, PM7, and DFTB3 models MAX errors stems from estimating the incorrect energy of CO, CF ($^4\Sigma$ state), and N≡C−C≡N species, respectively. Some DFT models, such as M062X/6-31G* and B3LYP-D3BJ/cc-pVTZ, gave large errors from aromatic nitrogen heterocycles, but BEBOP gives better results for these species. The validity of a BEBOP analysis of bonding in organic molecules with hetero-atoms must be determined on a case-by-case basis through comparison with *ab initio* energies.

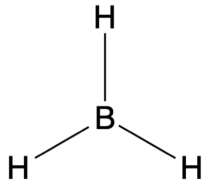
Inorganic Species

The wide range of bonding found in inorganic molecules presents a significant challenge for BEBOP and SQM methods. The species were selected to survey the possible bonding combinations. They include many species that have not been observed experimentally, but the consistent accuracy of the ROCBS-QB3 energies for the experimentally known species gives us confidence in these reference energies for the remainder. This test set includes a larger fraction of small molecules than the previous two sets, giving a deceptively small RMSE and MAD errors for BEBOP.

The 77 inorganic species include 31 examples from Table 2, which further reduces the BEBOP RMSE and MAD errors. The errors for the AM1, PM6, PM7, and DFTB3 models are much larger, reflecting that these models have difficulty modeling inorganic molecules. The DFT methods had the largest MAX error from B_3H_3 and B_2H_2. BEBOP energy for HBNH, however, deviates the most since ionic and charge transfer are critical in this structure. The BEBOP result for Be_4 is encouraging, as it suggests that bonding in metal clusters might be amenable to a BEBOP analysis.



	B1	B2	H1	H2	H3	H4	H5	H6
B1	89.99	-46.90	-59.20	-59.20	-121.82	-121.82	4.66	4.66
B2	-25.78	89.99	-59.20	-59.20	4.66	4.66	-121.82	-121.82
H1	-45.87	-45.87	0.00	11.08	2.28	2.28	2.28	2.28
H2	-45.87	-45.87	11.08	0.00	2.28	2.28	2.28	2.28
H3	-94.39	3.61	2.28	2.28	0.00	3.00	-0.16	0.22
H4	-94.39	3.61	2.28	2.28	3.00	0.00	0.22	-0.16
H5	3.61	-94.39	2.28	2.28	-0.16	0.22	0.00	3.00
H6	3.61	-94.39	2.28	2.28	0.22	-0.16	3.00	0.00



	B	H	H	H
B	81.90	-119.42	-119.42	-119.42
H	-92.12	0.00	4.11	4.11
H	-92.12	4.11	0.00	4.11
H	-92.12	4.11	4.11	0.00

Figure 8: BEBOP bond energy matrix depicting the net-bond energies (upper-diagonal elements), hybridization energies (diagonal elements), and gross bond energies (lower-diagonal elements) for diborane (with atom labels) and borane.

The 30.63 kcal/mol error in the BEBOP energy of diborane was initially disappointing, since we had hoped to achieve an improved understanding of the unusual three-center two-electron bonds. However, closer examination of the diborane and borane bond energy matrices (Figure 8) reveals terminal B-H net bond energies (−94.39 kcal/mol) almost unchanged from those in BH_3 (−92.12 kcal/mol). Combined with the bonding of each bridging hydrogen to each boron (−45.87 kcal/mol), the BEBOP model indicates a small (−8.20 kcal/mol) increase in the total directly bonded B-H bond energy. The energy of the B-B bond of diborane (−25.78 kcal/mol) would then account for the remainder of the $B_2H_6 \rightarrow 2BH_3$ dissociation energy ($D_{0,\text{exp.}} = 37.55$ kcal/mol). Unfortunately, the individually small, but numerous repulsive geminal interactions in diborane destroy the agreement with experiment. Nevertheless, the BEBOP description of the three-center two-electron bonds of diborane appears to be qualitatively correct and is not significantly altered by rescaling (by 1.052) to fit the ROCBS-QB3 energy (Figure 9).

Bond Dissociation Energies

The RMSE for 30 BEBOP bond dissociation energies (8.20 kcal/mol), $\Delta_{\text{BD}}E_0$ (Figure 10), is comparable to the error in the 192 total atomization energies (8.42 kcal/mol), D_0 (Table

	B1	B2	H1	H2	H3	H4	H5	H6
B1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B2	-27.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
H1	-48.26	-48.26	0.00	0.00	0.00	0.00	0.00	0.00
H2	-48.26	-48.26	11.65	0.00	0.00	0.00	0.00	0.00
H3	-99.30	3.80	2.40	2.40	0.00	0.00	0.00	0.00
H4	-99.30	3.80	2.40	2.40	3.16	0.00	0.00	0.00
H5	3.80	-99.30	2.40	2.40	-0.17	0.23	0.00	0.00
H6	3.80	-99.30	2.40	2.40	0.23	-0.17	3.16	0.00

Figure 9: Scaled BEBOP net bond energy matrix of the diborane molecule to fit ROCBS-QB3 data.

8). The majority of the hybrid DFT methods' absolute errors stems from underestimating the BDEs of C−H bonds in H−C≡C (∼20 kcal/mol) and the ortho- and para-positioned hydrogen from the phenyl radical (around or above 10 kcal/mol), but the BEBOP model has lower absolute deviation errors for the BDEs of these species (less than 10 kcal/mol). All semi-empirical models gave significant absolute deviation errors (above 20 kcal/mol) for C−H bonds present in the phenyl radical (except DFTB3) and the C=C bond of ethylene. Additionally, the AM1 and DFTB3 had difficulty estimating BDEs for N_2O and FOOF. Similarly, BEBOP had difficulty estimating the C=C BDE of ethylene (19.16 kcal/mol as shown in Figure 10), as the bonding of each CH_2 fragment is underestimated by 7.00 kcal/mol and the C=C bonding of ethylene is overestimated by 5.00 kcal/mol (Table 8). The failure of the BEBOP model to quantitatively describe the three-center two-electron bonds of diborane was noted previously. However, in most cases the BEBOP error is small enough to give us confidence in the BEBOP analysis.

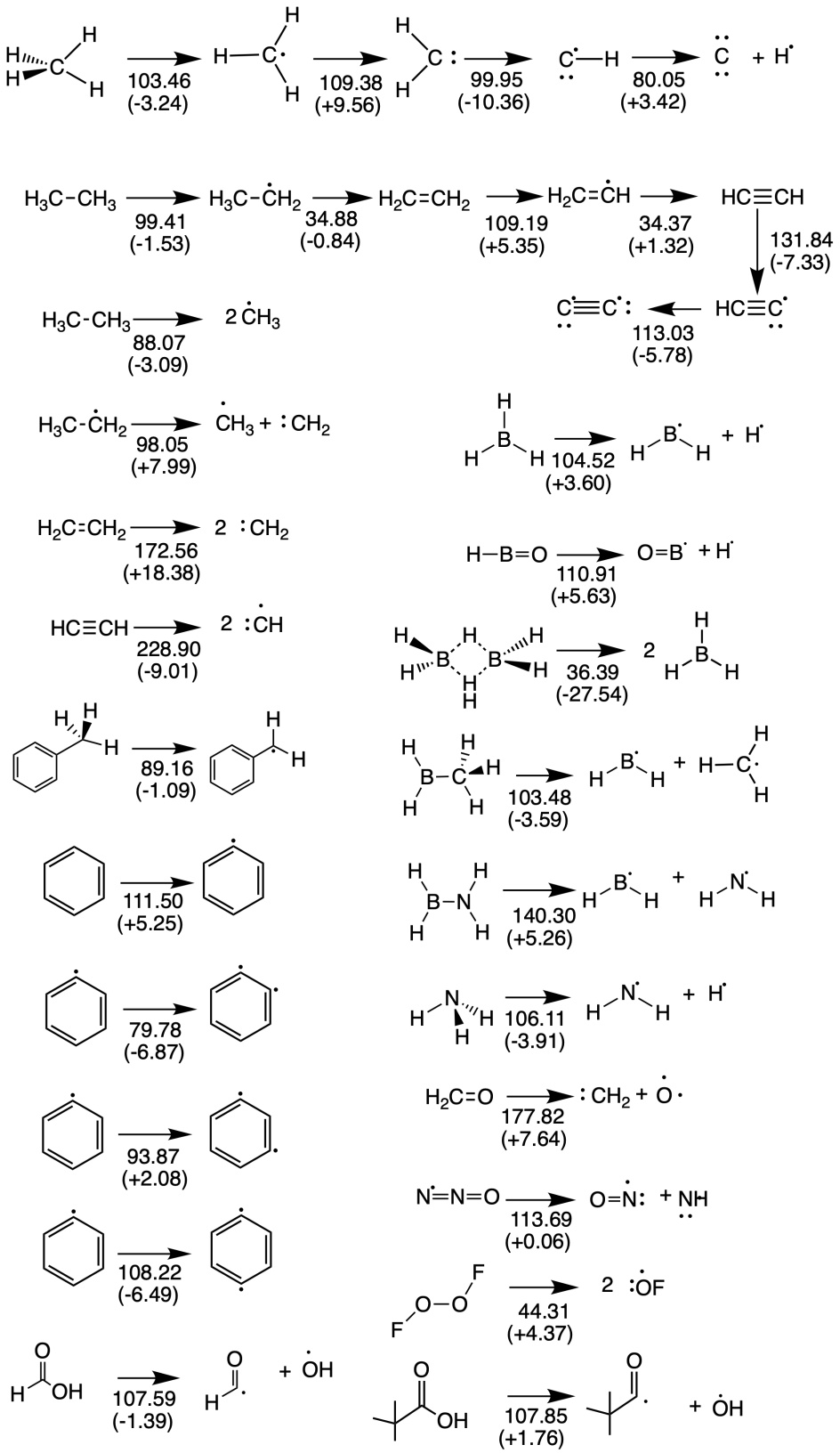


Figure 10: Deviation of BEBOP's BDE energy to experimental energies in kcal/mol from species used in our test set. All experimental D_0 were obtained from Ref. 36, except compounds containing B and $(CH_3)_3COOH$ which were obtained from ROCBS-QB3.

Table 9: Root mean squared errors (RMSE), mean absolute deviation (MAD), and maximum unsigned deviation (MAD), in kcal/mol, for the 30 bond dissociation energies at 0 K present in Figure 10.

Method	RMSE	MAX	MAD
thermochemical			
CBS-QB3//B3LYP/CBSB7	0.86	2.05	0.64
BEBOP//B3LYP/CBSB7	7.95	27.54	5.79
SQM			
PM7//PM7	13.17	27.82	11.04
DFTB3/3OB//DFTB3/3OB	15.73	34.28	11.77
PM6//PM6	15.99	37.84	13.00
AM1//AM1	19.45	42.80	15.86
DFT			
B3LYP/6-31G*//B3LYP/6-31G*	5.69	19.91	3.73
B3LYP-D3BJ/6-31G*//B3LYP/6-31G*	5.70	20.24	3.69
B3LYP-D3BJ/cc-pVTZ//B3LYP/6-31G*	5.85	24.07	3.78
B3LYP/CBSB7//B3LYP/CBSB7	5.88	23.14	3.78
B3LYP/cc-pVTZ//B3LYP/6-31G*	5.98	23.74	3.96
APF/6-31G*//APF/6-31G*	6.51	22.75	4.58
APFD/cc-pVTZ//APFD/6-31G*	6.67	26.07	4.58
APFD/6-31G*//APFD/6-31G*	6.56	22.86	4.63
APF/cc-pVTZ//APF/6-31G*	6.70	25.97	4.66
PBE0/6-31G*//PBE0/6-31G*	6.80	23.45	4.82
PBE0-D3BJ/6-31G*//PBE0/6-31G*	6.95	23.78	4.92
M062X/cc-pVTZ//M062X/6-31G*	7.14	29.14	3.90
M062X/6-31G*//M062X/6-31G*	7.13	25.42	4.38

The BEBOP model, in general, generated better D_{298} for the small alcohol molecules (Figure 11) than all SQMs and some popular hybrid DFT methods. As shown in Table 10, the RMSE (5.29 kcal/mol) is between B3LYP-D3BJ/6-31G* (5.02 kcal/mol) and APF/cc-pVTZ (5.60 kcal/mol). The MAD (3.33 kcal/mol) is also between these two methods. A deep analysis into the accuracy of D_{298} of each alcohol bond reveals where the current BEBOP model is limited.

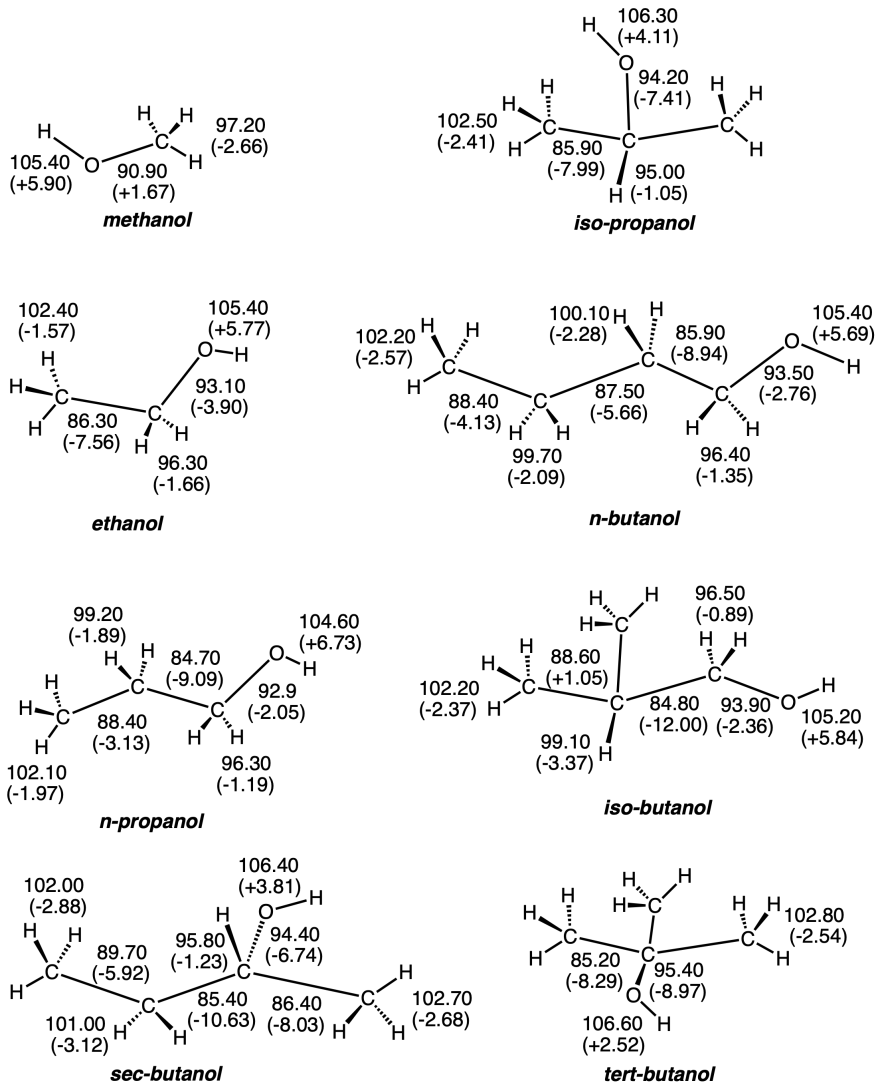


Figure 11: Deviation of BEBOP's and the reference[70] D_{298} energies for small alcohol molecules in kcal/mol.

The bulk of error from BEBOP comes from C−C, C−O, and O−H, but the C−C BDEs contributed the most error. For C−C bonds, our current model has no distinct Hückel-type energy and short-range repulsion parameters to describe σ-σ and π-π bonding. The C−O and O−H errors are not as large but significant for BEBOP relative to other hybrid DFT methods since the model is not modeling ionic bonding and charge transfer properly. However, BEBOP predicted the O−H D_{298} far more accurate than other DFT methods, except for M062X/cc-pVTZ and other SQM methods (not including PM7). In general, DFT methods were more accurate for nonpolar bonds (i.e., C−C and C−H) than polar bonds (i.e.,

O – H and C – O). A similar conclusion from Truhlar group's on BDEs of unsaturated methyl esters showed most exchange-correlation methods do not estimate polar BDEs to chemical accuracy.[71] Finally, the ROCBS-QB3 and ALFABET ML model, in general, gave accurate D_{298} for these species, with ROCBS-QB3 being the most accurate.

The ALFABET model is more competitive in giving accurate BDEs for species containing C,H,O, and H. However, the BEBOP model has strengths in generating QC-based bond energies essential for general chemistry applications. ML models can generate large numbers of optimizable parameters to extrapolate the energy from thousands of data. The BEBOP implementation is far more robust than ML models for the following reasons: (i) equations are physically rooted on the SQM formalism, (ii) the current parameters are physically-based and transferable across chemical space, (iii) the number of parameters is small, (iv) 36 reference species were used for parametrization, and (v) the model attempts to extrapolate the ROCBS-QB3 total electronic energy well for hydrocarbons (especially C – H bond in alcohol). Thus, the current BEBOP implementation and parameterization schemes of extrapolating the total energy are as simple but not robust as ML models because of current limitations.

The BEBOP model's current limitations for small molecules are the inability to model charge transfer, ionic bonding, and multivalent bond energies. For these reasons, we get large errors for O-H, C-O, and C-C BDEs. Thus, the current BEBOP model ensures that the reliability of the energetic properties of a species depends on the accuracy of atomization energy.

Applications

The differences in the energies of related species are of more chemical interest than the absolute energies. We will now introduce several tools that BEBOP seamlessly computes at low costs, such as resonance stabilization energies of aromatic species, relative energies of C_4H_6, strain energies of small ring compounds, and barriers for chemical reactions.

Table 10: Root mean squared error (RMSE), maximum absolute difference (MAX), and mean absolute deviation (MAD) for the bond dissociation energies D_{298} (in kcal/mol) for C−C, C−H, C−O, and O−H bonds of small alcohol isomers.

Method	RMSE					MAX					MAD				
	$D_{298,\text{all}}$	$D_{298,\text{C-C}}$	$D_{298,\text{C-H}}$	$D_{298,\text{C-O}}$	$D_{298,\text{O-H}}$	$D_{298,\text{all}}$	$D_{298,\text{C-C}}$	$D_{298,\text{C-H}}$	$D_{298,\text{C-O}}$	$D_{298,\text{O-H}}$	$D_{298,\text{all}}$	$D_{298,\text{C-C}}$	$D_{298,\text{C-H}}$	$D_{298,\text{C-O}}$	$D_{298,\text{O-H}}$
composite thermochemical															
CBS-QB3//B3LYP/CBSB7	1.17	1.02	1.20	1.38	1.05	2.11	1.75	2.11	1.97	1.65	1.04	0.35	0.55	0.23	0.40
BEBOP//B3LYP/CBSB7	5.29	8.03	2.20	5.20	5.22	12.00	12.00	3.37	8.97	6.73	3.33	1.84	0.60	2.73	1.17
ML Model															
ALFABET ML Model	2.11	0.71	2.60	1.90	2.43	4.00	1.40	4.0	2.40	3.40	1.63	0.55	0.63	0.41	0.54
SQM															
DFTB3/3OB//DFTB3/3OB	16.71	22.54	8.25	26.69	1.06	34.34	28.51	11.49	34.34	1.69	8.93	3.50	2.22	3.26	0.81
AM1//AM1	17.64	23.83	16.62	16.45	4.69	33.91	33.91	21.13	28.04	6.24	5.53	4.79	1.67	5.63	1.13
PM7//PM7	17.68	24.78	16.84	13.08	6.05	34.61	34.61	21.1	18.44	7.9	5.66	4.12	2.72	2.46	1.38
PM6//PM6	18.13	25.53	18.78	7.79	3.49	33.41	33.41	23.52	12.52	5.35	7.46	3.78	3.04	2.46	1.69
DFT															
M062X/cc-pVTZ//M062X/6-31G*	2.52	2.64	2.47	1.44	3.19	4.72	4.72	3.88	2.05	4.03	2.26	1.77	1.00	0.39	0.45
B3LYP-D3BJ/cc-pVTZ//B3LYP/6-31G*	4.52	4.09	3.71	4.54	6.56	7.49	7.42	5.39	6.49	7.49	1.61	1.81	1.31	0.77	0.46
APFD/cc-pVTZ//APFD/6-31G*	4.76	2.37	5.07	3.14	7.45	8.60	5.01	6.68	4.90	8.60	3.25	1.61	1.29	0.64	0.59
M062X/6-31G*//M062X/6-31G*	4.85	6.30	1.66	3.31	7.86	8.91	8.91	3.00	3.86	8.82	3.87	1.61	1.13	0.43	0.46
B3LYP-D3BJ/6-31G*//B3LYP/6-31G*	5.02	2.05	2.79	2.63	11.00	12.09	4.18	4.57	4.67	12.09	3.05	1.60	1.44	0.72	0.52
APF/cc-pVTZ//APF/6-31G*	5.60	4.42	5.65	4.44	7.82	8.72	8.4	7.73	7.03	8.72	3.56	1.85	1.37	1.16	0.49
APFD/6-31G*//APFD/6-31G*	5.62	2.94	3.93	1.63	11.75	13.05	5.69	5.76	3.31	13.05	4.17	1.48	1.42	0.65	0.63
B3LYP/6-31G*//B3LYP/6-31G*	5.78	3.57	3.44	4.31	11.64	12.62	7.14	5.52	7.18	12.62	2.85	1.94	1.54	1.18	0.48
PBE0-D3BJ/6-31G*//PBE0/6-31G*	5.86	3.44	4.23	1.44	12.00	13.12	6.41	6.27	3.27	13.12	3.96	1.58	1.44	0.77	0.55
APF/6-31G*//APF/6-31G*	5.92	2.27	4.50	2.92	12.13	13.18	4.55	6.63	5.44	13.18	3.98	1.75	1.48	1.04	0.51
B3LYP/cc-pVTZ//B3LYP/6-31G*	6.00	7.09	4.35	6.32	7.19	11.49	11.49	6.38	9	8.02	1.77	1.95	1.36	1.17	0.47
B3LYP/CBSB7//B3LYP/CBSB7	6.21	7.48	4.46	5.78	7.89	10.22	10.22	5.98	8.09	8.86	1.59	1.29	1.18	0.82	0.49
PBE0/6-31G*//PBE0/6-31G*	6.26	2.54	4.94	3.02	12.63	13.65	5	7.18	5.78	13.65	3.44	1.96	1.53	1.24	0.51

Resonance Energy

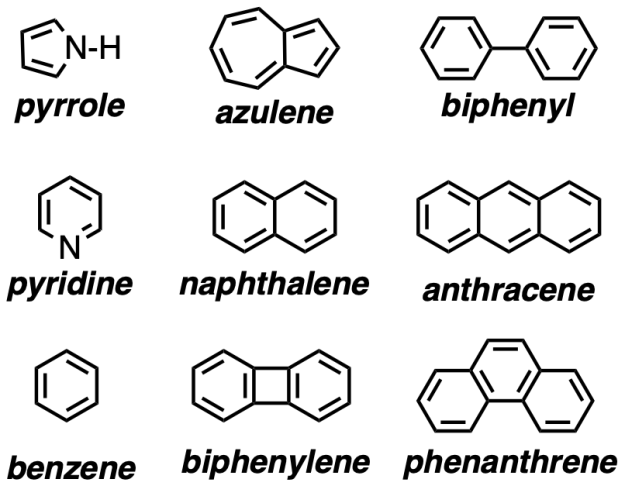


Figure 12: Aromatic structures used for computing BEBOP's resonance energies.

The extra stability of aromatic molecules is accurately reflected by the increased bond orders. The total atomization energies of a variety of aromatic species (Table 11) are thus well described by the BEBOP model (6.14 kcal/mol and 4.25 kcal/mol for RMSE and MAD, respectively). However, the real utility of the BEBOP model lies in the interpretation of these energies. The close association of the energies of aromatic species with bond orders invites an analysis of resonance stabilization energies based on these bond orders. The total BEBOP energy:

$$\begin{aligned} E_{\text{total}} &= \sum_{A\neq B}\left[\left(\sum_{\mu_A\nu_B}^{\text{AO}}\sum_{i}^{\text{MO}} C_{i\mu_A}C_{i\nu_B}S_{\mu_A\nu_B}\right)\beta_{A,B} + D_{AB}\exp(-\zeta_{AB}\left[R_{AB}-\frac{R_e(A,B)}{\sqrt{2}}\right])\right] \\ &+ \sum_{A}\left[n_{2\text{s}}(A)_{\text{ref}} - \sum_{i}^{\text{MO}}|C_{i2\text{s}_A}|^2\right]\Delta E(A)_{2\text{s}\rightarrow 2\text{p}} \\ &= \sum_{A\neq B}[P(A,B)\beta_{A,B} + E(A,B)_{\text{repulsion}}] + \sum_{A} E(A)_{\text{hybridization}} \end{aligned} \tag{21}$$

includes the bond-order-dependent covalent energy along with the short-range repulsion and hybridization terms, which vary with geometry and 2s population respectively, but are independent of the bond-order. The *resonance stabilization* of aromatic species can thus

be directly calculated as the difference between the total BEBOP energy and the BEBOP energy of a hypothetical species with the same geometry and hybridization, but with the bond orders associated with the Lewis structures (Figure 12):

$$\Delta E_{\text{resonance}} = \sum_{A \neq B} \left[\beta_{A,B} \left(P(A,B) - P(A,B)_{\text{reference}} \right) \right]. \tag{22}$$

For example, we subtract the reference bond-order for a C–C single bond (0.7874 for the central bond of 1, 3-butadiene twisted 90°) for each formal C–C single bond and the reference bond-order for a C=C double bond (1.1958 for the central bond of cis-2-butene) for each formal C=C double bond. This is identical to the definition used with simple Hückel theory but with reference π-bond-orders of 0 and 1, respectively.

Resonance stabilization energies derived from enthalpies of combustion vary considerably with the choice of reference species.[72–75] If we assign a reference C–C single bond order (0.8) that is intermediate between trans-1, 3-butadiene (0.8378) and the single bonds in cis-2-butene (0.7679), we obtain resonance stabilization energies in good agreement with the estimates of Dewar.[72–74] At the other extreme, if we assign a reference C–C single bond order (0.7760) that is intermediate between 1, 3-butadiene twisted 90° (0.7874) and cis-2-butene (0.7679), we obtain resonance stabilization energies in good agreement with the estimates of Allinger.[75] We prefer the middle ground, using the reference C–C single bond order from 1, 3-butadiene twisted 90° (0.7874). This species is obviously not an option as a reference for experimentally based estimates of resonance energies, but it provides a single bond between sp^2 hybridized carbons with no possibility of π-delocalization.

The "experimental" energies of species such as the hypothetical non-aromatic cyclohexatriene must necessarily refer to a relaxed geometry appropriate in the absence of delocalization, and cannot easily adjust for additional effects such as ring strain. However, our BEBOP estimate for biphenylene retains the low bond order (0.7363) for the strained C–C bonds of the four-membered ring, resulting in a somewhat larger estimate of the resonance

stabilization energy (44.6 kcal/mol vs. 39.0 kcal/mol).

Table 11: Comparison of *ab initio* and BEBOP total atomization energies (kcal/mol) for aromatic species, and comparison of BEBOP resonance energies with experiment. The "Non-Aromatic" energy substitutes bond orders[a] consistent with the structures drawn.

Species	CBS-QB3	Aromatic BEBOP	Non-Aromatic[a]	Resonance Energy				
				$\Delta_c H_{298}$	BEBOP[a]/C-C BO=			$\Delta_c H_{298}$
				Dewar[b]	0.8000	0.7874	0.7760	Allinger[c]
pyrrole	-1019.61	-1033.69	-1012.99		17.8	20.7	23.2	21.2
pyridine	-1182.11	-1183.61	-1168.41		9.5	15.2	20.3	22.7
benzene	-1303.74	-1307.94	-1279.74	20	19.7	28.2	36	36.1
azulene	-2034.47	-2042.8	-2020.2	4.2	5.6	22.6	38.1	33.3
naphthalene	-2068.97	-2070.21	-2022.31	30.5	30.8	47.9	63.3	61.4
biphenylene	-2344.7	-2339.6	-2333.7		-16.9	5.9	26.5	39.0[d]
biphenyl	-2503.73	-2503.79	-2450.59		33.3	53.2	71.3	70.7
anthracene	-2831.3	-2829.66	-2765.26	36.9	38.8	64.4	87.6	84.3
phenanthrene	-2836.88	-2834.67	-2764.67	44.6	44.3	70.0	93.1	92.7

[a] C−C → 0.7874; C=C → 1.1958; C−N· → 0.6941; C=N → 1.0699; =C−N: → 0.7644.
[b] Table 5.4, p177, in Ref. 73.
[c] Table 11.3, p252 and Table 13.1, p284, in Ref. 75.
[d] Table 9-2, p244, in Ref. 76.

The choice of appropriate reference bond orders is important to obtain meaningful results. We employ the reference C–N bond order (0.6941) from $H_2C{=}CH{-}N{=}CH_2$ twisted 90°, and the reference C=N bond order (1.0699) from *cis*-$CH_3CH{=}NCH_3$ for pyridine, but use the reference C–N: bond order (0.7644) from planar vinyl amine, $H_2CCH{-}NH_2$, for pyrrole. In general, one should compare to the most structurally similar acyclic species available. Although our choice of reference is ultimately still subjective, we believe the BEBOP "non-aromatic" resonance stabilization energies in Table 11 are a realistic measure of the *specific effects of delocalization* in these species. The BEBOP model will accommodate whatever reference one finds appropriate.

Strain Energy

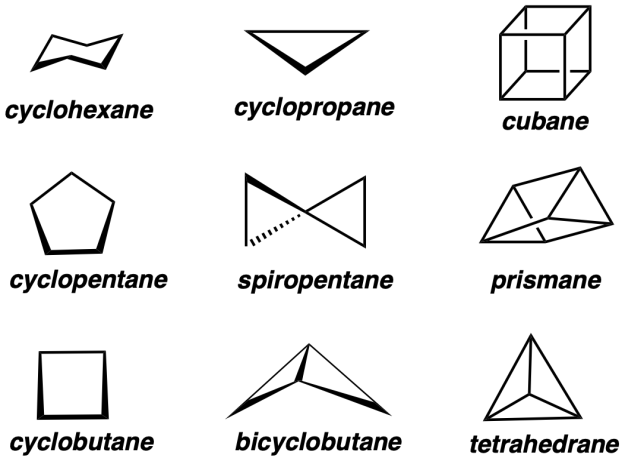


Figure 13: Strained structures used for computing BEBOP's strain energies.

The reduced stability of molecules with ring strain is again reflected by the (now reduced) bond orders. The total atomization energies of a variety of strained species (Figure 13) are thus well described by the BEBOP model (3.56 and 2.81 kcal/mol RMSE and MAD, respectively). However, our focus now is on the interpretation of these energies.

Quantitative estimates of strain energies once again raise the issue of selecting a suitable reference. The ring strain in the simple cycloalkanes can be readily estimated by assigning one sixth of the energy of cyclohexane (i.e., 1657.00/6 = 276.17 kcal/mol) to a strain-free cyclic methylene, CH_2, group. The difference between three times this value and the calculated energy for cyclopropane is then the strain energy (Table 12).

We obtain virtually the same strain energies for cyclopentane (6.13 kcal/mol), cyclobutane (25.82 kcal/mol), and cyclopropane (26.95 kcal/mol), if we employ experimental heats of formation.[36]

Table 12: Comparison of *ab initio* and BEBOP total atomization energies (kcal/mol) for strained species. The "Unstrained" energy substitutes bond orders consistent with the absence of ring strain.

Species	CBS-QB3	BEBOP		ΔE	Strain Energy
		Strained	Unstrained		
cyclohexane	-1657.00	-1656.87		0.00	0
cyclopentane	-1374.69	-1377.42	-1385.13	7.71	6.13
cyclobutane	-1078.83	-1082.44	-1116.19	33.75	25.82
cyclopropane	-801.28	-808.80	-839.91	31.11	26.95
spiropentane	-1211.12	-1212.26	-1272.54	60.28	
bicyclobutane	-930.19	-924.94	-987.24	62.30	
cubane	-1617.24	-1616.82	-1826.56	209.74	
prismane	-1189.09	-1187.22	-1314.31	127.09	
tetrahedrane	-753.42	-751.38	-938.17	186.79	

With the cycloalkanes as a starting point, we have developed a procedure analogous to our treatment of resonance stabilization for a BEBOP analysis of ring strain. We treat the C–C bond order in cyclohexane (0.7336) as a reference for a strain-free cyclic C–C bond, except in three-membered rings. The $(1,3)$ antibonding interaction between alternant carbons in cyclohexane (-0.0465) must be included in the reference bond order for both adjacent carbons in a 3-membered ring, since these directly bonded carbons are also geminal in a 3-membered ring. We therefore employ 0.6406 as a reference for a strain free C–C bond in a 3-membered ring. The resulting "unstrained" energies in Table 12 give energy changes, $\Delta E = E_{\text{strained}} - E_{\text{unstrained}}$ in reasonable agreement with the energetically derived strain energies for the cycloalkanes. They also provide a plausible estimate of the ring strain in species such as spiropentane and tetrahedrane.

Valence Isomers

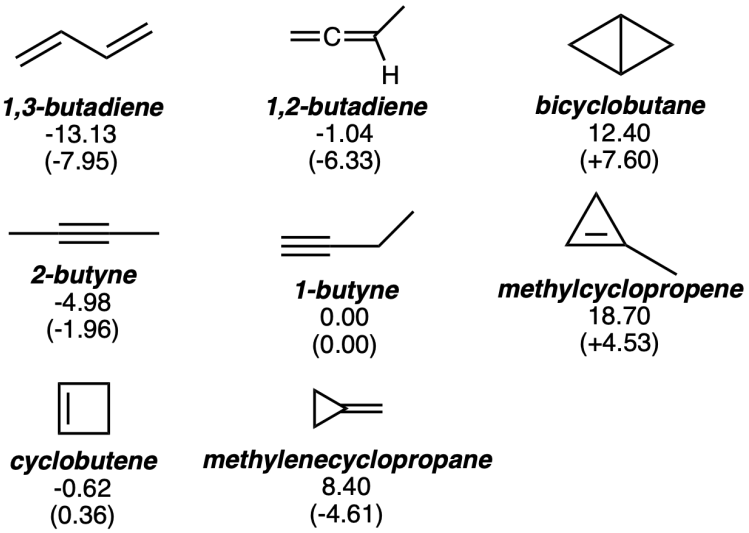


Figure 14: Deviation of BEBOP's relative energies (shown in parentheses) of the C_4H_6 isomers compared to experimental energies in kcal/mol. All experimental relative energies were obtained from Ref. 36, except for the experimental relative energies of methylenecyclopropane, bicyclobutane, and methylcyclopropene obtained from Ref. 77.

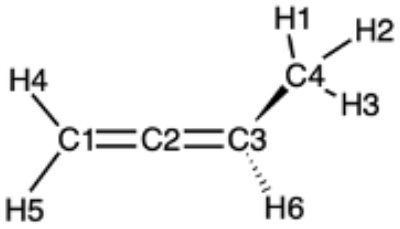
The BEBOP model facilitates a detailed analysis of the relative energies of valence isomers (Figure 14). The relative energies of eight C_4H_6 isomers are qualitatively correct, but the range is about 1/3 too large (Table 13).

The ordering is qualitatively correct except 1, 2-butadiene, and the largest MAX error is from 1, 3-butadiene. In general, DFT methods gave lower errors than BEBOP and SQMs. However, SQMs do rank the relative energies of these species qualitatively except for cyclobutene and methylcyclopropene (for AM1, PM6, and PM7 models only). The RMSE and MAD errors of BEBOP (5.17 kcal/mol and 4.41 kcal/mol, respectively) are comparable to those of aromatic and strained species. On one hand, the RMSE is intermediate between APFD/cc-pVTZ (4.64 kcal/mol) and DFTB3 (5.48 kcal/mol). On the other hand, the MAD error is intermediate between DFTB3 (3.37 kcal/mol) and PM6 (5.68 kcal/mol). The most interesting comparisons are the structures with the same number of single, double, and triple bonds, such as 1, 2-butadiene vs. 1, 3-butadiene. These are energy differences that are accurately reflected by the bond orders.

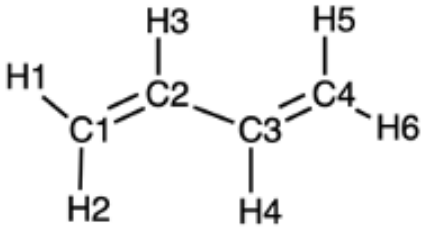
Table 13: Root mean squared error (RMSE), maximum absolute difference (MAX), and mean absolute deviation (MAD) for the relative energies (in kcal/mol) of C_4H_6 isomers.

Methods	RMSE	MAX	MAD
composite thermochemical			
CBS-QB3//B3LYP/CBSB7	1.07	1.93	0.76
BEBOP//B3LYP/CBSB7	5.08	7.95	4.17
SQM			
DFTB3/3OB//DFTB3/3OB	5.36	10.86	3.27
PM6//PM6	8.05	17.95	5.40
PM7//PM7	9.66	23.32	6.18
AM1//AM1	11.07	28.20	6.38
DFT			
B3LYP/CBSB7//B3LYP/CBSB7	2.49	4.31	1.93
B3LYP-D3BJ/cc-pVTZ//B3LYP/6-31G*	2.69	4.07	2.03
B3LYP/cc-pVTZ//B3LYP/6-31G*	2.73	4.15	2.40
M062X/cc-pVTZ//M062X/6-31G*	3.02	5.15	2.14
B3LYP/6-31G*//B3LYP/6-31G*	3.57	5.56	1.77
B3LYP-D3BJ/6-31G*//B3LYP/6-31G*	3.61	5.68	1.70
M062X/6-31G*//M062X/6-31G*	4.52	7.13	2.75
APF/cc-pVTZ//APF/6-31G*	4.68	7.83	2.00
APFD/cc-pVTZ//APFD/6-31G*	4.75	7.77	2.29
APF/6-31G*//APF/6-31G*	6.40	9.58	2.99
APFD/6-31G*//APFD/6-31G*	6.52	9.54	3.29
PBE0/6-31G*//PBE0/6-31G*	6.76	9.96	3.30
PBE0-D3BJ/6-31G*//PBE0/6-31G*	6.84	10.08	3.42

Thus far, we have focused on the total BEBOP energies. The relative differences are more reliably obtained from the *ab initio* calculations than using BEBOP. The more interesting results of the BEBOP analysis come from the partitioning of the energy over the individual bonds.



	C1	C2	C3	C4	H1	H2	H3	H4	H5	H6
C1	82.62	-238.04	4.99	0.12	-0.06	0.00	0.00	-127.63	-127.63	0.10
C2	-148.26	89.94	-238.37	12.54	1.03	-0.17	-0.17	9.80	9.80	11.05
C3	3.29	-148.17	84.72	-161.55	7.95	8.45	8.45	0.09	0.09	-127.07
C4	0.08	8.10	-110.40	75.13	-125.08	-124.35	-124.35	0.00	0.00	8.51
H1	-0.05	0.82	6.59	-106.80	0.00	6.46	6.46	0.00	0.00	-0.50
H2	0.00	-0.13	7.01	-106.18	6.46	0.00	7.11	0.00	0.00	0.37
H3	0.00	-0.13	7.01	-106.18	6.46	7.11	0.00	0.00	0.00	0.37
H4	-106.03	7.76	0.07	0.00	0.00	0.00	0.00	0.00	6.83	0.00
H5	-106.03	7.76	0.07	0.00	0.00	0.00	0.00	6.83	0.00	0.00
H6	0.09	8.76	-105.41	7.26	-0.50	0.37	0.37	0.00	0.00	0.00



	C1	C2	C3	C4	H1	H2	H3	H4	H5	H6
C1	79.39	-233.66	13.11	-0.28	-128.83	-127.69	9.88	1.01	-0.01	0.01
C2	-154.71	82.90	-175.19	13.11	8.77	9.40	-127.63	9.88	1.44	-0.49
C3	8.68	-116.43	82.90	-233.66	-0.49	1.44	9.88	-127.63	9.40	8.77
C4	-0.19	8.68	-154.71	79.39	0.01	-0.01	1.01	9.88	-127.69	-128.83
H1	-106.90	7.30	-0.41	0.01	0.00	7.24	1.32	0.00	0.00	0.00
H2	-105.96	7.82	1.20	-0.01	7.24	0.00	-0.60	-0.33	-0.00	0.00
H3	8.20	-106.23	8.23	0.84	1.32	-0.60	0.00	-0.50	-0.33	0.00
H4	0.84	8.23	-106.23	8.20	0.00	-0.33	-0.50	0.00	-0.60	1.32
H5	-0.01	1.20	7.82	-105.96	0.00	0.00	-0.33	-0.60	0.00	7.24
H6	0.01	-0.41	7.30	-106.90	0.00	0.00	0.00	1.32	7.24	0.00

Figure 15: BEBOP bond energy matrix depicting the gross bond energy (upper-diagonal elements), the hybridization energy (diagonal elements), and net bond energies (lower-diagonal elements) in kcal/mol for 1,2-butadiene and *trans*-1,3-butadiene.

The total gross bonding and anti-bonding of the C∼C bond energies favor $1,2$-butadiene (-620.31 kcal/mol) over trans-$1,3$-butadiene (-616.57 kcal/mol) by 3.74 kcal/mol, but carbon 2 of $1,2$-butadiene has no hydrogen to share the cost of hybridization, so the net C∼C bond energies (-395.36 kcal/mol and -408.68 kcal/mol, respectively) favor trans-$1,3$-

butadiene by over 13.32 kcal/mol. A similar hybridization effect was seen in the comparison of the ethynyl radical with acetylene (*vide supra*). The BEBOP analysis correctly identifies trans-1, 3-butadiene as the more stable isomer and supports the conventional wisdom that conjugation enhances the strength of the C–C single bond in 1, 3-butadiene. Twisting 1, 3-butadiene by 90° about this bond to break the conjugation raises both the CBS-QB3 energy and the BEBOP energy by 5.80 kcal/mol and results in C=C and C–C BEBOP net bond energies of −155.50 and −110.40 kcal/mol respectively, directly verifying the source for half of the extra stability of the 1, 3 isomer. The remainder can be attributed to the larger hybridization energy (89.94 kcal/mol) for the *sp* hybridized carbon in the 1, 2 isomer, which propagates into reduced C=C BEBOP net bond energies in the 1, 2 isomer. The gross bond energies indicate that the extra energy required for *sp* hybridization is the problem, rather than a reduced ability of the orbitals of the 1, 2 isomer to form the C=C bonds. The final contribution to the energy difference is the reduced C$\cdots$C 1, 3 antibonding interaction in the linear 1, 2 isomer, favoring this isomer by over 4.99 kcal/mol. The BEBOP model provides a detailed quantitative and objective version of generally accepted qualitative concepts of chemical bonding in the isomers of butadiene.

Qualitative Chemical Reaction Mechanism Analyses

In general, single reference quantum chemistry methods are not recommended for analyzing structures away from stationary points on a potential energy surface. Since orbital populations used by BEBOP can be from either single reference unrestricted or restricted (open-shell) HF calculations, this BEBOP model in its present form is also generally not recommended for such applications. However, we wanted to test the interpretability of BEBOP energy decompositions along reaction pathways where a single reference HF calculation would be acceptable. A detailed BEBOP analysis of the barriers for a sequence of hydrogen abstraction reactions provides a clear example of the increased level of understanding that the method offers. We employ the standard BEBOP analysis using UHF to describe

bond breaking appropriately and UCBS-QB33 reference energies along the intrinsic reaction pathway for UMP2/6-311G(2d,d,p) optimized geometries.

Beginning with the most elementary example, the reaction of H_2 with a hydrogen atom (Figure 16):

$$H_2 + H\cdot \rightarrow H\cdots H\cdots H \rightarrow H + H_2$$

we see that the formation of the new H–H bond (BE_{23}) strongly overlaps the disappearance of the old H–H bond (BE_{12}), so that the change in the sum of these bond energies ($\Delta E_{12}+\Delta E_{23}$) is very small. We could reasonably describe the mechanism as the bond moving from the $1,2$ to the $2,3$ hydrogen pair. The small barrier (9.1 kcal/mol) for this reaction results from the $1,3$-antibonding interaction of the unpaired electron originating on the hydrogen atom.

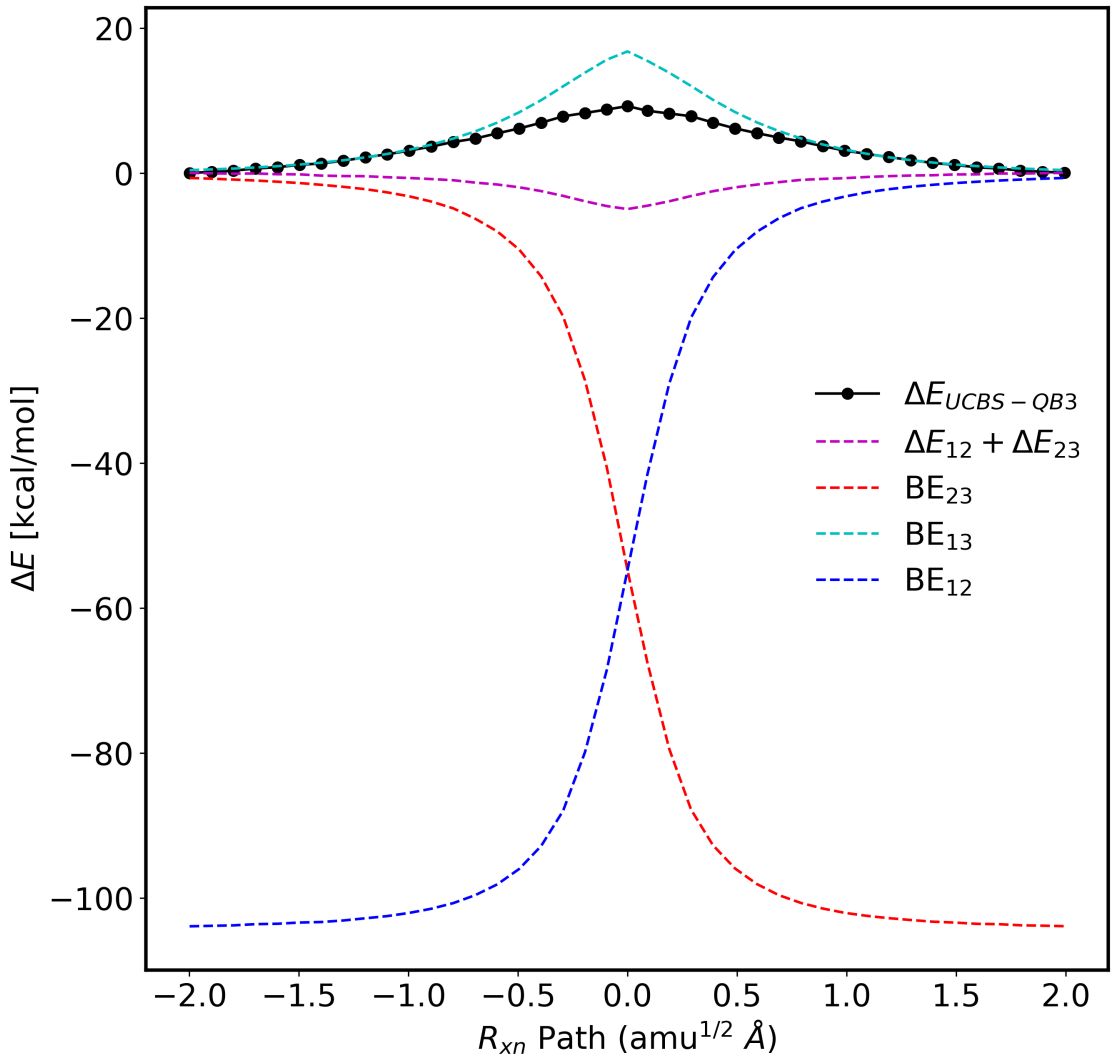


Figure 16: The variation of the BEBOP net bond energies and the ROCBS-QB3 total energy along the reaction path for the $H_2 + H\cdot \rightarrow H\cdot + H_2$ reaction.

The reaction of methane with a methyl radical:

$$H_3C{-}H + \cdot CH_3 \rightarrow H_3C \cdots H \cdots CH_3 \rightarrow H_3C\cdot + H{-}CH_3$$

very closely resembles the H_3 reaction, but the C···C 1, 3-antibonding interaction is a bit more repulsive ($\Delta E_0^{\ddagger} \approx 17$ kcal/mol). Once again, the formation of the new bond strongly overlaps the disappearance of the old bond (Figure 17).

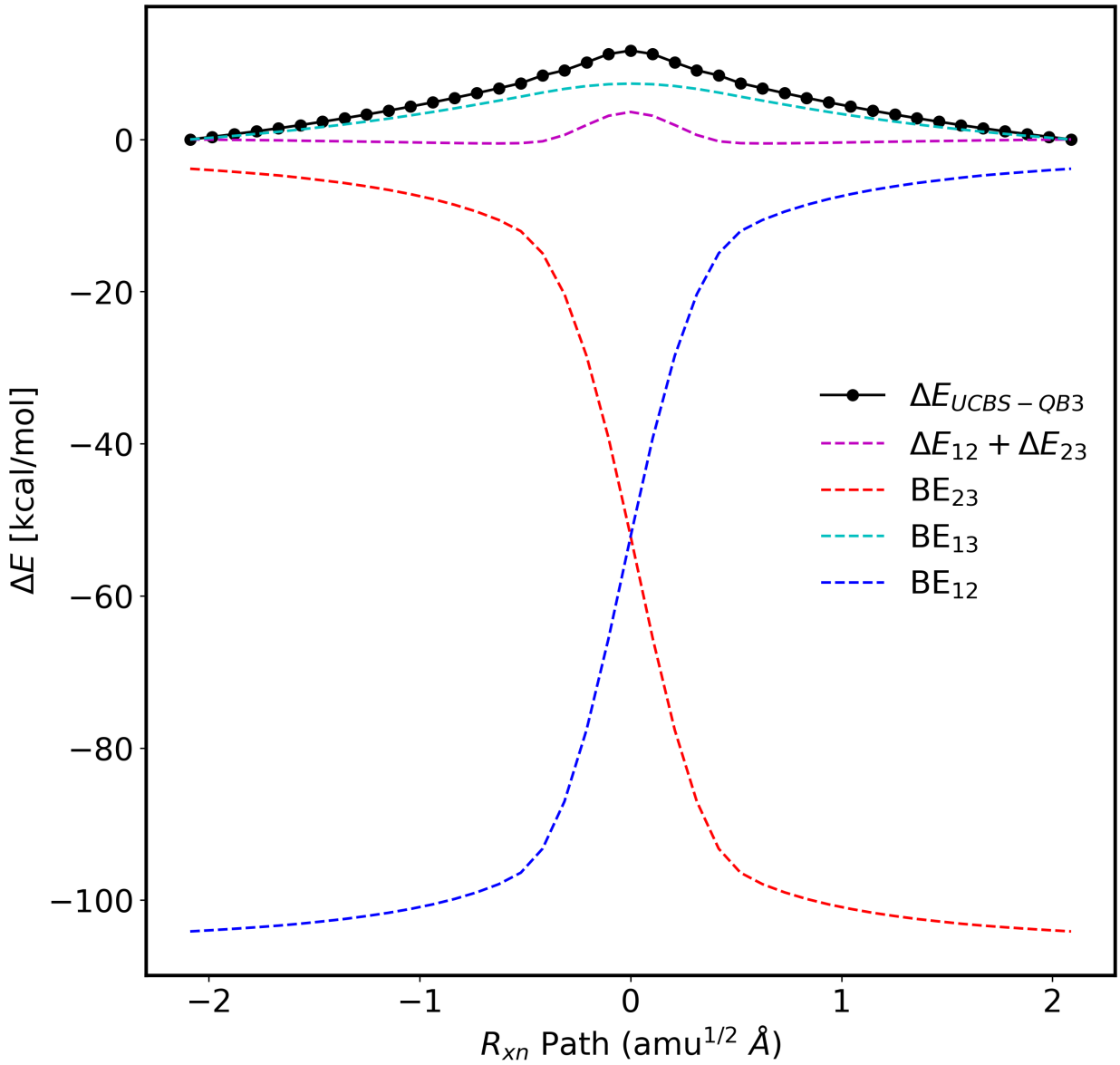
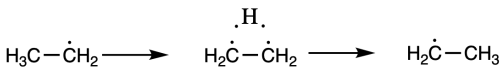


Figure 17: The variation of the BEBOP net bond energies and the ROCBS-QB3 total energy along the reaction path for the $H_3C-H + \cdot CH_3 \rightarrow H_3C\cdot + H-CH_3$ reaction.

The isomerization of ethyl radical:



presents a new contribution to the barrier for hydrogen abstraction. The directly bonded carbons cannot form hybrid orbitals directed at an acute angle to the C−C bond as required to maintain bonding to the migrating hydrogen. This provides an additional ∼ 20 kcal/mol,

raising the barrier to $\sim$ 40 kcal/mol (Figure 18). Note that the C$\cdots$C 1,3-antibonding interaction is now superimposed on the C–C σ-bond and thus exhibits as a reduction in the ethyl radical C–C bond energy from $\sim$ 105 kcal/mol in the reactant to $\sim$ 80 kcal/mol in the transition state. The closer proximity of these directly bonded carbons increases this 1,3-antibonding relative to the CH_4 + CH_3 reaction.

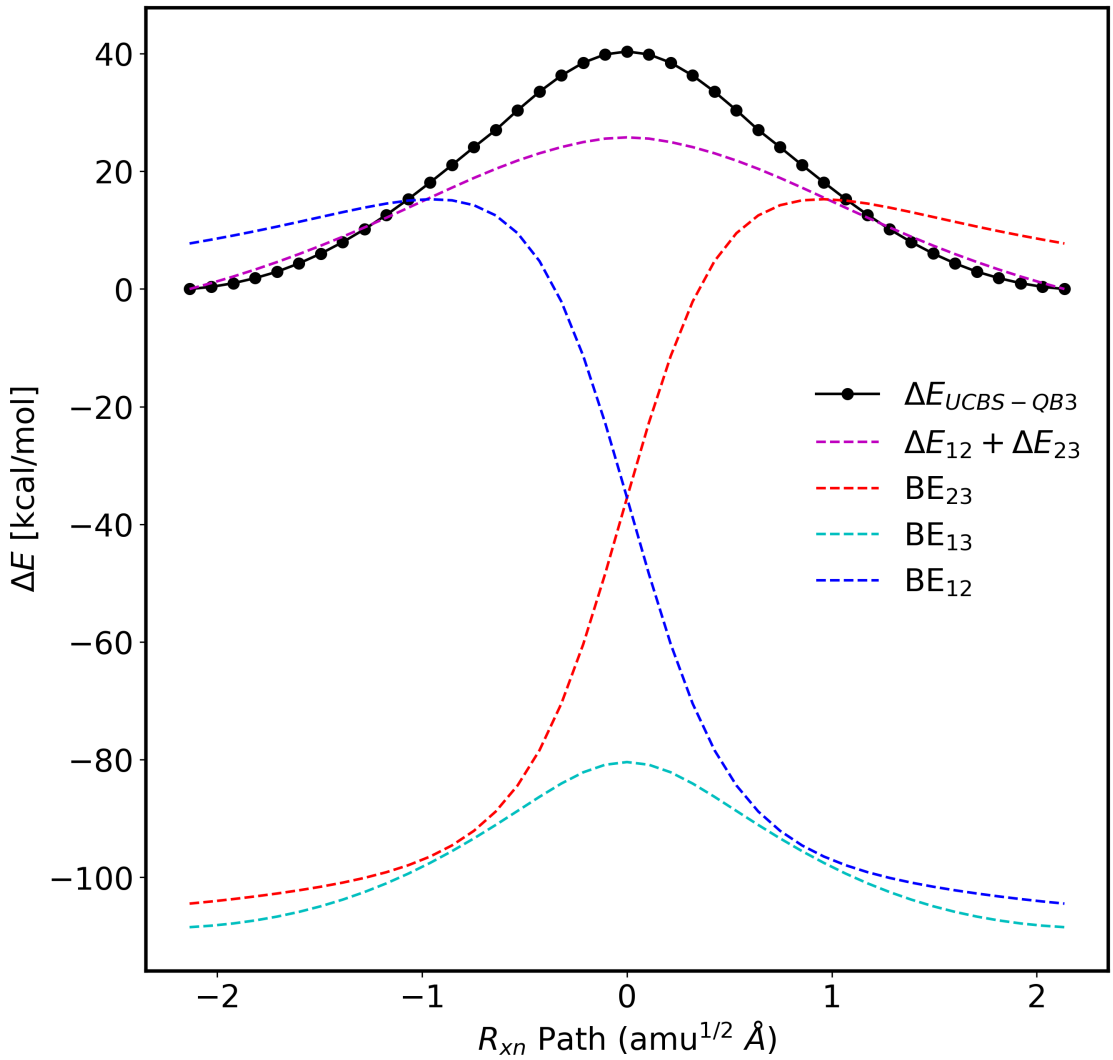
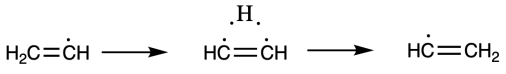


Figure 18: The variation of the BEBOP net bond energies and the ROCBS-QB3 total energy along the reaction path for the $H_3C\text{-}\dot{C}H_2 \rightarrow H_2\dot{C}\text{–}CH_3$ reaction.

The isomerization of the vinyl radical (Figure 19):



closely follows the behavior of the ethyl radical, but the C=C bond energy is of course larger than the C−C bond energy of the ethyl radical. The reduced C=C bond length increases the barrier to ∼ 47 kcal/mol. The C···H···C 1, 3 antibonding (BE_{13}) and the inability to maintain bonding to the migrating hydrogen ($\Delta E_{12}+\Delta E_{23}$) both contribute to this substantial barrier.

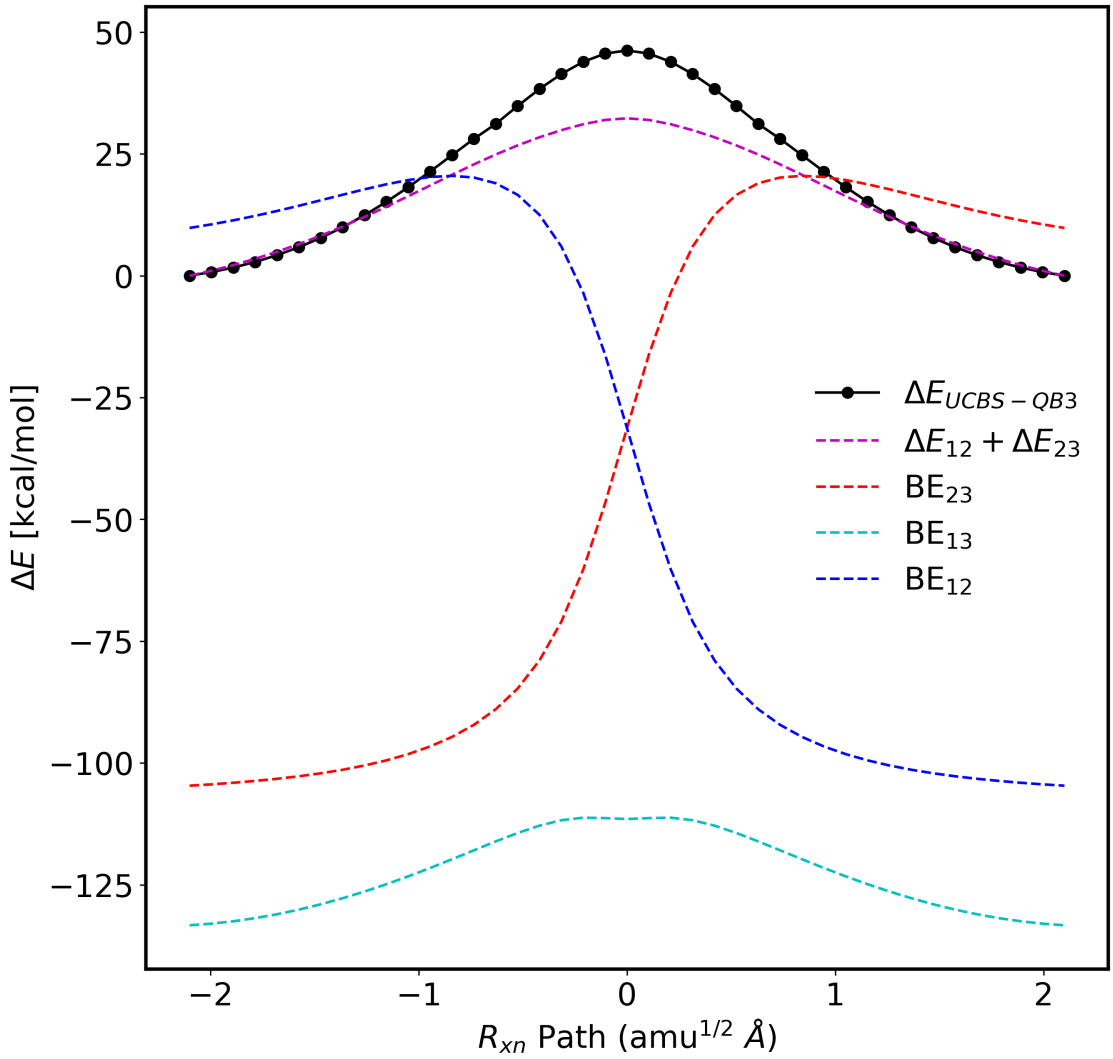
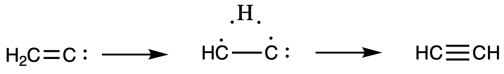


Figure 19: The variation of the BEBOP net bond energies and the ROCBS-QB3 total energy along the reaction path for the $H_2C{=}\,\dot{C}H \rightarrow H\dot{C}{=}CH_2$ reaction.

The isomerization of vinylidene (Figure 20):



presents a qualitatively different potential energy surface with a very small (∼ 2 kcal/mol)

barrier. The contribution from the inability to bind the migrating hydrogen ($\Delta E_{12} + \Delta E_{23}$) tightly is comparable, but the 1,3-antibonding interaction is now superimposed on the newly forming C–C π-bond. The C–C bond energy (BE_{13}) increases to 187 kcal/mol in the product, acetylene. If we partition this bond energy into the σ- and π-components, we find the π-component, $BE\pi_{13}$, monotonically increasing in strength, while the σ-component, $BE\sigma_{13}$, presents a reaction barrier (Figure 20) similar to those in Figure 18 and Figure 19.

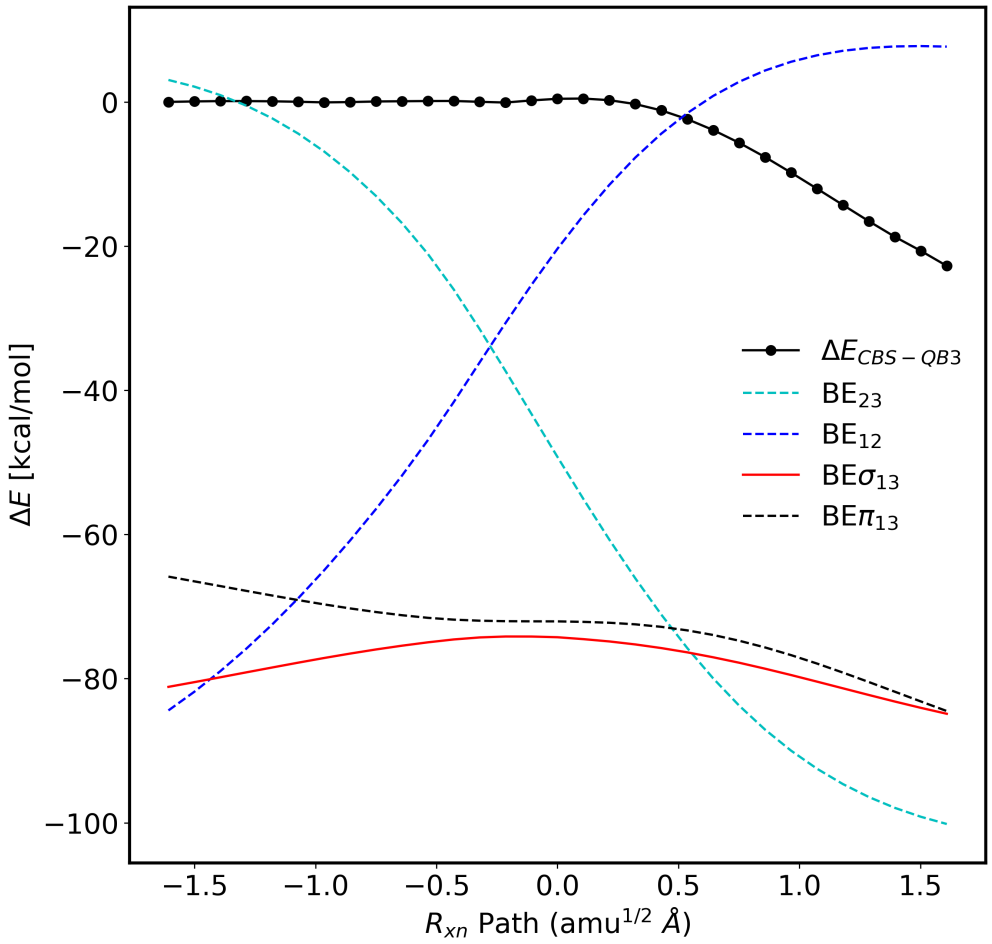


Figure 20: The variation of the BEBOP net bond energies and the CBS-QB3 total energy along the reaction path for the $H_2C{=}C{:} \rightarrow HC{\equiv}CH$ reaction.

The vinylidene isomerization was a subject of considerable interest a few years ago.[78] This strongly exothermic ($\Delta E_0 = -44$ kcal/mol) reaction has a very low barrier ($\Delta E_0^{\ddagger} = 2$ kcal/mol) and a transition state structure resembles the product as much as the starting material. The energy profile along the reaction path for the total energy (i.e. CBS-QB3

in Figure 21) apparently violates the Hammond postulate.[79] However, we can interpret the vinylidene isomerization as the superimposition of two distinct processes: the transfer of the migrating hydrogen, and the conversion of the vinylidene lone pair to the new C∼C π-bond in the acetylene product. The barrier height (i.e., the peak in $\Delta E_{12} + \Delta E_{23} + E_{\sigma_{13}}$ in Figure 21) and transition state structure for the hydrogen migration are consistent with their counterparts for isomerization of the ethyl (Figure 18) and vinyl (Figure 19) radicals. The vinylidene reactant is the transition state for the formation of the new C∼C π-bond ($\Delta E_{\pi_{13}}$ in Figure 21). Each of these processes individually satisfies the Hammond postulate. We would expect the Hammond postulate to describe individual reaction processes more consistently than composite reactions. These are the same conclusions we reached previously,[78] but the BEBOP analysis is free from subjective decisions and is far less complicated than our earlier analysis.

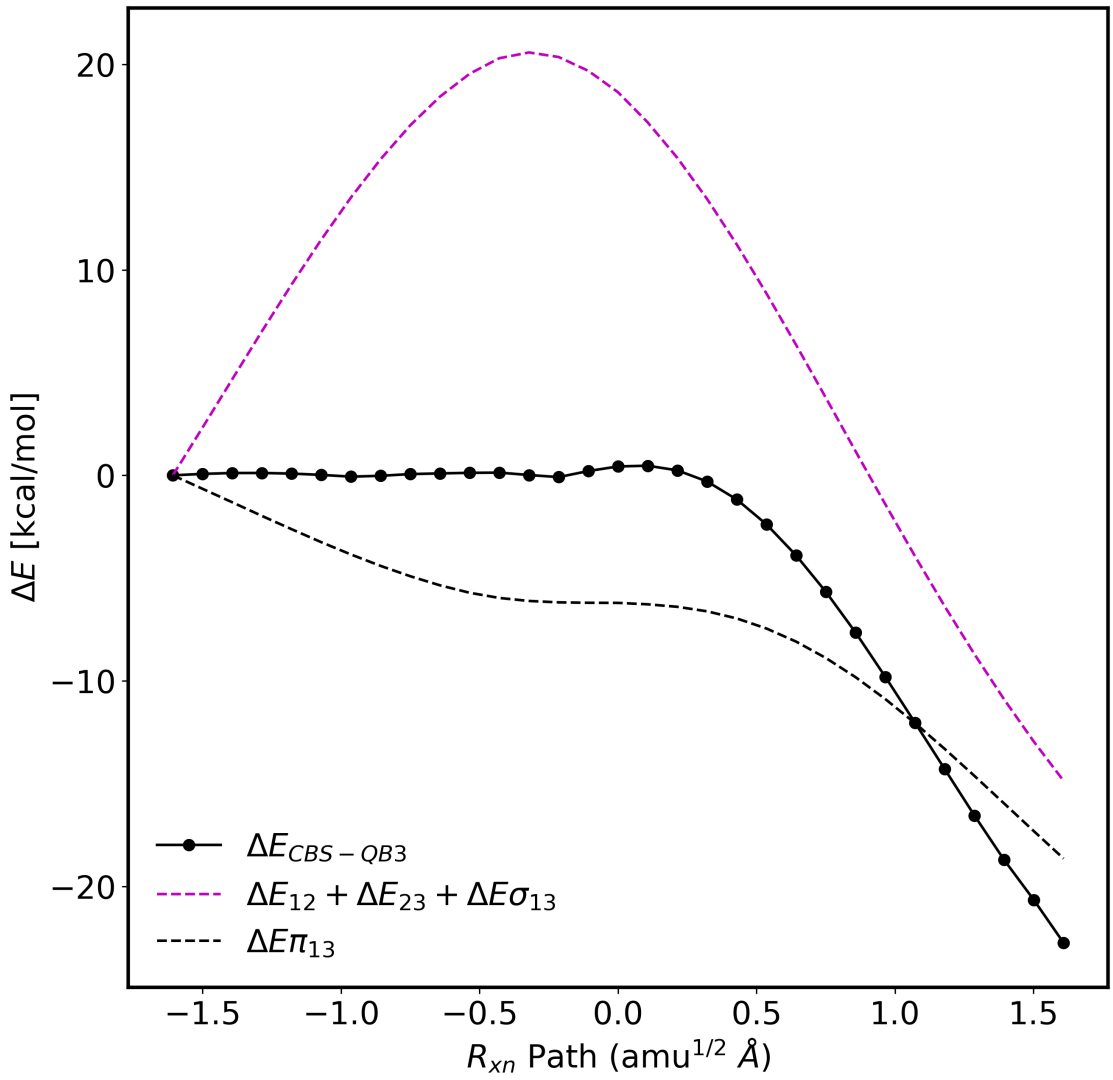


Figure 21: The variation of the BEBOP net bond energies and the CBS-QB3 total energy along the reaction path for the $H_2C{=}C{:} \rightarrow HC{\equiv}CH$ reaction.

Limitations

As authors are prone to do, we have emphasized the strengths of our model. However, it is also important to appreciate the limitations of any method.

1. Ionic species are not properly described by the purely covalent bond energy analysis of the BEBOP model. We obtain the correct bond energy for species such as LiF in our calibration study through the artifact of an unrealistically large value for β_{LiF}, but

this can provide no insight into the nature of the bonding in ionic species.

2. We must be cautious in applications to species other than hydrocarbons. The current BEBOP model overestimates the strength of each =C−F bond by $\sim$ 10 kcal/mol, and often underestimates the strength of C=O bonds by a comparable error. These effects combine to produce a BEBOP error of more than 30 kcal/mol for the C−F bond dissociation energy of $F_2C{=}O$.

3. The BEBOP model recovers only $\sim$ 25% of the dissociation energy of diborane, B_2H_6, to form two molecules of borane, BH_3. This is disappointing in view of the interesting three-center two-electron bonds in diborane. However, the qualitative BEBOP description of these bonds appears to be correct.

We strongly caution against the application of the current BEBOP analysis to any inorganic species unless a reliable *ab initio* energy is available to verify the BEBOP energy.

Conclusions

We have developed BEBOP, a bond energies from bond orders and populations model, for the partition of ab initio energies among the bonds of a molecule. The necessary parameters have been determined for all bonds involving H, Li, Be, B, C, N, O, and F atoms. The BEBOP model reproduces the atomization energies of 56 hydrocarbon molecules and radicals to within ±4.80 kcal/mol RMSE. The model provides reliable and qualitative analyses of resonance stabilization energies, ring strain energies, and variations in bond dissociation energies in hydrocarbons. The BEBOP analysis of transition state barriers for chemical reactions offers new qualitative insights about the nature of these barriers. The reliability of the model deteriorates as we introduce hetero-atoms (±10.40 kcal/mol RMSE for 60 species) and for applications to inorganic chemistry (±8.77 kcal/mol RMSE for 76 species), but the potential for useful qualitative interpretations of ab initio energies is very broad. Our future

work with the BEBOP model will also extend parameters for second row elements. We hope to include additional expressions for charge transfer and ionic effects that are expected to improve the performance for inorganic and hetero-organic species. We also plan to include an empirical dispersion correction to model long range electron correlations.

Acknowledgement

BZ acknowledges support from the K. Leroy Irvis Fellowship through the University of Pittsburgh Swanson School of Engineering PITT STRIVE program. JAK acknowledges support from the NSF (CBET-1653392). We thank the University of Pittsburgh Center for Research Computing for computing time and technical support. The graduate studies of SVT. were supported by United Technologies Corporation. This work was partially supported by the National Center for Supercomputing Applications under Grant #CTS-060029N to PRW and utilized the "cobalt" computer system. The work at Wesleyan was partially supported by Gaussian, Inc.

Supporting Information Available

The following files are available free of charge.

- BEBOP_Tables.xlsx: Spreadsheet containing benchmarks for BEBOP and other computational chemistry methods
- BEBOP codes: available in `https://github.com/keithgroup`

References

(1) Pauling, L. The nature of the chemical bond. IV. The energy of single bonds and the relative electronegativity of atoms. *J. Am. Chem. Soc.* **1932**, *53*, 3570–3582, DOI:

https://doi.org/10.1021/ja01348a011.

(2) Mulliken, R. S. Electronic population analysis on LCAO-MO molecular wave functions. I. *J. Chem. Phys.* **1955**, *23*, 1833–1840, DOI: 10.1063/1.1740588.

(3) Mulliken, R. S. Criteria for the construction of good self-consistent-field molecular orbital wave functions, and the significance of LCAO-MO population analysis. *J. Chem. Phys.* **1962**, *36*, 3428–3439, DOI: 10.1063/1.1732476.

(4) Glendening, E. D.; Landis, C. R.; Weinhold, F. Natural bond orbital methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 1–42, DOI: 10.1002/wcms.51.

(5) Weinhold, F. Natural bond orbital analysis: A critical overview of relationships to alternative bonding perspectives. *J. Comput. Chem.* **2012**, *33*, 2363–2379, DOI: 10.1002/jcc.23060.

(6) Glendening, E. D.; Landis, C. R.; Weinhold, F. NBO 6.0: Natural bond orbital analysis program. *J. Comput. Chem.* **2013**, *34*, 1429–1437, DOI: 10.1002/jcc.23266.

(7) Brenner, D. W. Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. *Phys. Rev. B* **1990**, *42*, 9458–9471, DOI: 10.1103/PhysRevB.42.9458.

(8) Pettifor, D. G.; Oleinik, I. I. Analytic bond-order potentials beyond Tersoff-Brenner. I. Theory. *Phys. Rev. B* **1999**, *59*, 8487–8499, DOI: 10.1103/PhysRevB.59.8487.

(9) Tersoff, J. New empirical approach for the structure and energy of covalent systems. *Phys. Rev. B* **1988**, *37*, 6991–7000, DOI: 10.1103/PhysRevB.37.6991.

(10) Van Duin, A. C.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A reactive force field for hydrocarbons. *J. Phys. Chem. A* **2001**, *105*, 9396–9409, DOI: 10.1021/jp004368u.

(11) Christensen, A. S.; Sirumalla, S. K.; Qiao, Z.; O'Connor, M. B.; Smith, D. G. A.; Ding, F.; Bygrave, P. J.; Anandkumar, A.; Welborn, M.; Manby, F. R.; Miller, T. F. OrbNet Denali: A machine learning potential for biological and organic chemistry with semi-empirical cost and DFT accuracy. *J. Chem. Phys.* **2021**, *155*, 204103, DOI: 10.1063/5.0061990.

(12) Diefenbach, A.; Bickelhaupt, F. M.; Frenking, G. The Nature of the Transition Metal Carbonyl Bond and the Question about the Valence Orbitals of Transition Metals. A Bond-Energy Decomposition Analysis of TM(CO)6q (TMq= Hf2-, Ta-, W, Re+, Os2+, Ir3+). *J. Am. Chem. Soc.* **2000**, *122*, 6449–6458, DOI: 10.1021/ja000663g.

(13) Su, P.; Li, H. Energy decomposition analysis of covalent bonds and intermolecular interactions. *J. Chem. Phys.* **2009**, *131*, 014102, DOI: 10.1063/1.3159673.

(14) Levine, D. S.; Head-Gordon, M. Energy decomposition analysis of single bonds within Kohn–Sham density functional theory. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 12649–12656, DOI: 10.1073/pnas.1715763114.

(15) Thirman, J.; Engelage, E.; Huber, S. M.; Head-Gordon, M. Characterizing the interplay of Pauli repulsion, electrostatics, dispersion and charge transfer in halogen bonding with energy decomposition analysis. *Phys. Chem. Chem. Phys.* **2018**, *20*, 905–915, DOI: 10.1039/C7CP06959F.

(16) Mao, Y.; Loipersberger, M.; Horn, P. R.; Das, A.; Demerdash, O.; Levine, D. S.; Prasad Veccham, S.; Head-Gordon, T.; Head-Gordon, M. From Intermolecular Interaction Energies and Observable Shifts to Component Contributions and Back Again: A Tale of Variational Energy Decomposition Analysis. *Annu. Rev. Phys. Chem.* **2021**, *72*, 641–666, DOI: 10.1146/annurev-physchem-090419-115149.

(17) Hoffmann, R. An extended Hückel theory. I. Hydrocarbons. *J. Chem. Phys.* **1963**, *39*, 1397–1412, DOI: 10.1063/1.1734456.

(18) Thiel, W. Semiempirical quantum-chemical methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 145–157, DOI: 10.1002/wcms.1161.

(19) Christensen, A. S.; Kromann, J. C.; Jensen, J. H.; Cui, Q. Intermolecular interactions in the condensed phase: Evaluation of semi-empirical quantum mechanical methods. *J. Chem. Phys.* **2017**, *147*, DOI: 10.1063/1.4985605.

(20) González, R.; Suárez, C. F.; Bohórquez, H. J.; Patarroyo, M. A.; Patarroyo, M. E. Semi-empirical quantum evaluation of peptide – MHC class II binding. *Chem. Phys. Lett.* **2017**, *668*, 29–34, DOI: 10.1016/j.cplett.2016.12.015.

(21) Hunt, P.; Hosseini-Gerami, L.; Chrien, T.; Plante, J.; Ponting, D. J.; Segall, M. Predicting pKa Using a Combination of Semi-Empirical Quantum Mechanics and Radial Basis Function Methods. *J. Chem. Inf. Model.* **2020**, *60*, 2989–2997, DOI: 10.1021/acs.jcim.0c00105.

(22) Li, P.; Jia, X.; Pan, X.; Shao, Y.; Mei, Y. Accelerated Computation of Free Energy Profile at ab Initio Quantum Mechanical/Molecular Mechanics Accuracy via a Semi-Empirical Reference Potential. I. Weighted Thermodynamics Perturbation. *J. Chem. Theory Comput.* **2018**, *14*, 5583–5596, DOI: 10.1021/acs.jctc.8b00571.

(23) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. A complete basis set model chemistry. VII. Use of the minimum population localization method. *J. Chem. Phys.* **2000**, *112*, 6532–6542, DOI: 10.1063/1.481224.

(24) Wood, G. P.; Radom, L.; Petersson, G. A.; Barnes, E. C.; Frisch, M. J.; Montgomery, J. A. A restricted-open-shell complete-basis-set model chemistry. *J. Chem. Phys.* **2006**, *125*, 94106–94122, DOI: 10.1063/1.2335438.

(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.;

Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian16, Revision C.01. 2016.

(26) Lennard-Jones, J. E. Cohesion. *Proc. Phys. Soc* **1931**, *43*, 462–482.

(27) London, F. W. The general theory of molecular forces. *J. Chem. Soc. Faraday Trans.* **1937**, *33*, 8–26.

(28) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. An effective fragment method for modeling solvent effects in quantum mechanical calculations. *J. Chem. Phys.* **1996**, *105*, 1968–1986, DOI: 10.1063/1.472045.

(29) Ben-Nun, M.; Martinez, T. J. Direct evaluation of the Pauli repulsion energy using 'classical' wavefunctions in hybrid quantum/classical potential energy surfaces. *Chem. Phys. Lett.* **1998**, *290*, 289–295, DOI: https://doi.org/10.1016/S0009-2614(98)00486-2.

(30) Chalmet, S.; Ruiz-Lopez, M. F. New approaches to the description of short-range repulsion interactions in hybrid quantum/classical systems. *Chem. Phys. Lett.* **2000**, *329*, 154–159.

(31) Grimme, S. A General Quantum Mechanically Derived Force Field (QMDFF) for

Molecules and Condensed Phase Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 4497–4514, DOI: 10.1021/ct500573f.

(32) Wilson, A. L.; Popelier, P. L. A. Exponential Relationships Capturing Atomistic Short-Range Repulsion from the Interacting Quantum Atoms (IQA) Method. *J. Phys Chem. A* **2016**, *120*, 9647–9659, DOI: 10.1021/acs.jpca.6b10295.

(33) Van Vleet, M. J.; Misquitta, A. J.; Stone, A. J.; Schmidt, J. R. Beyond Born–Mayer: Improved Models for Short-Range Repulsion in ab Initio Force Fields. *J. Chem. Theory Comput.* **2016**, *12*, 3851–3870, DOI: 10.1021/acs.jctc.6b00209.

(34) Ruscic, B.; Pinzon, R. E.; Morton, M. L.; von Laszevski, G.; Bittner, S. J.; Nijsure, S. G.; Amin, K. A.; Minkoff, M.; Wagner, A. F. Introduction to Active Thermochemical Tables: Several "Key" Enthalpies of Formation Revisited. *J. Phys. Chem. A* **2004**, *108*, 9979–9997, DOI: 10.1021/jp047912y.

(35) Ruscic, B.; Pinzon, R. E.; Laszewski, G. v.; Kodeboyina, D.; Burcat, A.; Leahy, D.; Montoy, D.; Wagner, A. F. Active Thermochemical Tables: thermochemistry for the 21st century. *J. Phys. Conf. Ser.* **2005**, *16*, 561–570, DOI: 10.1088/1742-6596/16/1/078.

(36) Ruscic, B.; Bross, D. Active Thermochemical Tables (ATcT) values based on ver. 1.122r of the Thermochemical Network. 2021; ATcT.anl.gov.

(37) Roothaan, C. C. A Study of Two-Center Integrals Useful in Calculations on Molecular Structure. I. *J. Chem. Phys.* **1951**, *19*, 1445–1458, DOI: 10.1063/1.1748100.

(38) Hehre, W. J.; Stewart, R. F.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *J. Chem. Phys.* **1969**, *51*, 2657–2664, DOI: 10.1063/1.1672392.

(39) Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. Self-Consistent Molecular Orbital Methods. IV. Use of Gaussian Expansions of Slater-Type Orbitals. Extension to Second-Row Molecules. *J. Chem. Phys.* **1970**, *52*, 2769–2773, DOI: 10.1063/1.1673374.

(40) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776, DOI: 10.1063/1.477422.

(41) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652, DOI: 10.1063/1.464913.

(42) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Physical Review Letters* **1996**, *77*, 3865–3868, DOI: 10.1103/PhysRevLett.77.3865.

(43) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]. *Phys. Rev. Lett.* **1997**, *78*, 1396–1396, DOI: 10.1103/PhysRevLett.78.1396.

(44) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170, DOI: 10.1063/1.478522.

(45) Austin, A.; Petersson, G. A.; Frisch, M. J.; Dobek, F. J.; Scalmani, G.; Throssell, K. A density functional with spherical atom dispersion terms. *J. Chem. Theory Comput.* **2012**, *8*, 4989–5007, DOI: 10.1021/ct300778e.

(46) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class

functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241, DOI: 10.1007/s00214-007-0310-x.

(47) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-consistent molecular-orbital methods. IX. An extended gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **1971**, *54*, 720–723, DOI: 10.1063/1.1674902.

(48) Hehre, W. J.; Ditchfield, K.; Pople, J. A. Self-consistent molecular orbital methods. XII. Further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261, DOI: 10.1063/1.1677527.

(49) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theoret. chim. Acta (Berl.)* **1973**, *28*, 213–222.

(50) Dill, J. D.; Pople, J. A. Self-consistent molecular orbital methods. XV. Extended Gaussian-type basis sets for lithium, beryllium, and boron. 1975.

(51) Binkley, J. S.; Pople, J. A. Self-consistent molecular orbital methods. XIX. Split-valence Gaussian-type basis sets for beryllium. *J. Chem. Phys.* **1977**, *66*, 879–880, DOI: 10.1063/1.433929.

(52) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023, DOI: 10.1063/1.456153.

(53) Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning, T. H.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg. *Theor. Chem. Acc.* **2011**, *128*, 69–82, DOI: 10.1007/s00214-010-0764-0.

(54) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate <i>ab ini-

tio</i> parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104, DOI: 10.1063/1.3382344.

(55) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465, DOI: 10.1002/jcc.21759.

(56) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc* **1985**, *107*, 3902–3909, DOI: 10.1021/ja00299a024.

(57) Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **2007**, *13*, 1173–1213, DOI: 10.1007/s00894-007-0233-4.

(58) Stewart, J. J. P. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *J. Mol. Model.* **2013**, *19*, 1–32, DOI: 10.1007/s00894-012-1667-x.

(59) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method: Third-Order Expansion of the Density Functional Theory Total Energy and Introduction of a Modified Effective Coulomb Interaction. *J. Phys. Chem. A* **2007**, *111*, 10861–10873, DOI: 10.1021/jp074167r.

(60) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948, DOI: 10.1021/ct100684s.

(61) Stewart, J. MOPAC2016. 2016; http://openmopac.net/.

(62) Hourahine, B.; Aradi, B.; Blum, V.; Bonafé, F.; Buccheri, A.; Camacho, C.; Cevallos, C.; Deshaye, M. Y.; Dumitrică, T.; Dominguez, A.; Ehlert, S.; Elstner, M.; van der

Heide, T.; Hermann, J.; Irle, S.; Kranz, J. J.; Köhler, C.; Kowalczyk, T.; Kubař, T.; Lee, I. S.; Lutsker, V.; Maurer, R. J.; Min, S. K.; Mitchell, I.; Negre, C.; Niehaus, T. A.; Niklasson, A. M. N.; Page, A. J.; Pecchia, A.; Penazzi, G.; Persson, M. P.; Řezáč, J.; Sánchez, C. G.; Sternberg, M.; Stöhr, M.; Stuckenberg, F.; Tkatchenko, A.; Yu, V. W.-z.; Frauenheim, T. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **2020**, *152*, 124101, DOI: 10.1063/1.5143190.

(63) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338–354, DOI: 10.1021/ct300849w.

(64) Kubillus, M.; Kubař, T.; Gaus, M.; Řezáč, J.; Elstner, M. Parameterization of the DFTB3 Method for Br, Ca, Cl, F, I, K, and Na in Organic and Biological Systems. *J. Chem. Theory Comput.* **2015**, *11*, 332–342, DOI: 10.1021/ct5009137.

(65) St. John, P. C.; Guan, Y.; Kim, Y.; Kim, S.; Paton, R. S. Prediction of organic homolytic bond dissociation enthalpies at near chemical accuracy with sub-second computational cost. *Nature Commun.* **2020**, *11*, 2328, DOI: 10.1038/s41467-020-16201-z.

(66) St. John, P. C.; Guan, Y.; Kim, Y.; Etz, B. D.; Kim, S.; Paton, R. S. Quantum chemical calculations for over 200,000 organic radical species and 40,000 associated closed-shell molecules. *Sci. Data* **2020**, *7*, 244, DOI: 10.1038/s41597-020-00588-x.

(67) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297, DOI: 10.1039/b508541a.

(68) Gdanitz, R. J.; Ahlrichs, R. The averaged coupled-pair functional (ACPF): A size-extensive modification of MRCI(SD). *Chem. Phys. Lett.* **1988**, *143*, 413–420, DOI: 10.1016/0009-2614(88)87388-3.

(69) Gdanitz, R. J. A new version of the multireference averaged coupled-pair functional (MR-ACPF-2). *Int. J. Quantum Chem.* **2001**, *85*, 281–300, DOI: 10.1002/qua.10019.

(70) Oyeyemi, V. B.; Keith, J. A.; Carter, E. A. Trends in Bond Dissociation Energies of Alcohols and Aldehydes Computed with Multireference Averaged Coupled-Pair Functional Theory. *J. Phys. Chem. A* **2014**, *118*, 3039–3050, DOI: 10.1021/jp501636r.

(71) Li, X.; Xu, X.; You, X.; Truhlar, D. G. Benchmark Calculations for Bond Dissociation Enthalpies of Unsaturated Methyl Esters and the Bond Dissociation Enthalpies of Methyl Linolenate. *J. Phys. Chem. A* **2016**, *120*, 4025–4036, DOI: 10.1021/acs.jpca.6b02600.

(72) Dewar, M. J. S.; Gleicher, G. J. Physical and Inorganic Chemistry Ground States of Conjugated Molecules. II. Allowance for Molecular Geometry. *J. Am. Chem. Soc.* **1965**, *87*, 685–692.

(73) Dewar, M. J. *The molecular orbital theory of organic chemistry*, 1st ed.; McGraw-Hill Book Co.: New York, 1969; Chapter 5.

(74) S Dewar, M. J.; de Llano, C. Physical and Inorganic Chemistry Ground States of Conjugated Molecules. XI. Improved Treatment of Hydrocarbons. *J. Am. Chem. Soc.* **1969**, *91*, 789–795.

(75) Allinger, N.; Cava, M.; DeJongh, D.; Johnson, C.; LeBel, N.; Stevens, C. *Organic chemistry*; Worth Publishers: New York, 1971.

(76) Roberts, J. D.; Caserio, M. C. *Basic Principles of Organic Chemistry*; W. A. Benjamin, Inc.: Menlo Park, California, 1964.

(77) Pedley, J. *Thermochemical data and structures of organic compounds (Trc data series)*, 1st ed.; CRC Press: College Station, Texas, 1994; Vol. 1.

(78) Petersson, G. A.; Tensfeldt, T. G.; Montgomery, J. A. Vinylidene and the Hammond Postulate. *J. Am. Chem. Soc.* **1992**, *114*, 6133–6138, DOI: 10.1021/ja00041a034.

(79) Hammond, G. S. A Correlation of Reaction Rates. *J. Am. Chem. Soc.* **1955**, *77*, 334–338.